第2章

・スタージェスの公式

$$k \fallingdotseq 1+\log_2 n$$

k：階級数　n：観測値の数

・4尺度　　名義(名目)尺度，順序尺度，間隔尺度，比尺度

○代表値 … 分布の位置を示す指標

・算術平均

$$\bar{x} = \frac{x_1+x_2+\cdots+x_n}{n}$$

$$\bar{x} = a_1x_1+a_2x_2+\cdots+a_nx_n$$　　a_i：相対度数

※最初と最後の階級値が定まらない時

- 最初，最後の値を異常値とし，無視する
- 階級値を適当に決めて与える

・幾何平均　… 地価上昇率

$$x_G = \sqrt[n]{x_1\cdot x_2\cdot\cdots\cdot x_n}$$

・調和平均　… バスの平均時速

$$\frac{1}{x_H} = \frac{1}{n}\left(\frac{1}{x_1}+\cdots+\frac{1}{x_n}\right)$$

・メディアン，中位数，中央数

観測値に偏りがある場合，観測値を小さいものから順に並べて真ん中の値をメディアンとする。

他に百分位点というものがある。

・モード，最頻値

分布の峰に対応する値。ただし，峰が二つ以上ある場合は有効な代表値とは言えない。

・有用性

右に歪んだ分布　　Mean > Median > Mode

左　〃　　Mean < Median < Mode

・ミッドレンジ　$M = \frac{1}{2}\{\max(x_1, x_2, \cdots x_n)+\min(x_1, x_2, \cdots x_n)\}$

No. Date

○散らばりの尺度 … 分布の形状を示す指標

・レンジ

$$R = \max(x_1, x_2 \cdots x_n) - \min(x_1, x_2 \cdots x_n)$$

・四分位偏差

$$Q = \frac{1}{2}(Q_3 - Q_1)$$ Q_3：第3四分位点 Q_1：第1四分位点

・平均偏差

$$d = \frac{1}{n}\{|x_1 - \bar{x}| + |x_2 - \bar{x}| + \cdots + |x_n - \bar{x}|\}$$

・分散と標準偏差

$$S^2 = \frac{1}{n}\{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2\}$$ … 分散 $\frac{1}{n}\sum(x_i - \bar{x})^2$

$x_1^2 + \cdots + x_n^2 - n\bar{x}^2$ $1/n$

$$S = \sqrt{S^2}$$ … 標準偏差

・変動係数

$$C.V. = \frac{S_x}{\bar{x}}$$

・$z_i = ax_i + b$，$\bar{z} = a\bar{x} + b$ とした時 $S_z^2 = a^2 S_x$，$S_z = |a| S_x$

証明

$$S_z^2 = \sum_{i=1}^{n}(z_i - \bar{z})^2 \times \frac{1}{n}$$

$$= \sum_{i=1}^{n}\{(ax_i + b) - (a\bar{x} + b)\}^2 \times \frac{1}{n}$$

$$= a^2 \sum_{i=1}^{n}(x_i - \bar{x})^2 \times \frac{1}{n}$$

$$= a^2 \cdot S_x^2$$

・データxの標準化，標準得点

$$z_i = \frac{x_i - \bar{x}}{S_x}$$ … z得点

$$T_i = 10z_i + 50$$ … T得点，偏差値得点

? nが小さいとき $\mu \pm \sigma$
nが大きいとき $\mu \pm 2\sigma$

Date . . No.

第3章

・相関

変数に区別を設けず、対等に見る見方や方法

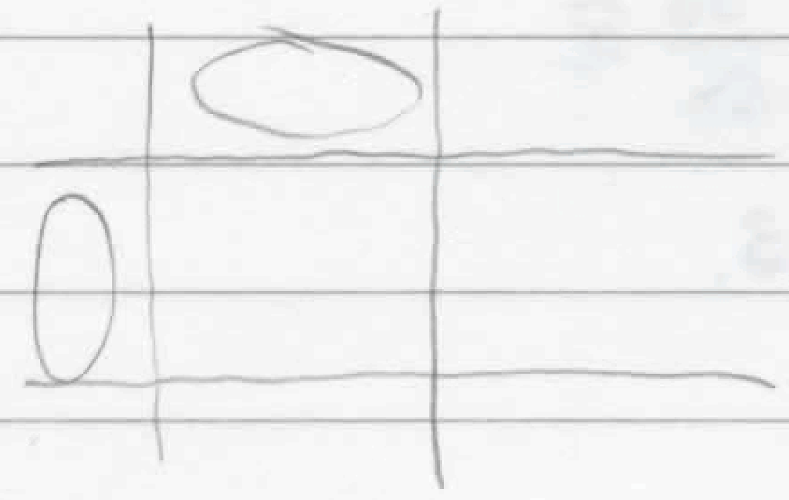
・回帰

特定の変数からある関数を見ること。

・分割表、クロス表

縦方向に書いてある変数を表側

横方向に書いてある変数を表頭と呼ぶ



・相関係数

相関の程度を表す指標

・ピアソンの積率相関係数

データが $(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)$ で与えられた時、相関係数は

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2}\sqrt{\sum (y_i - \bar{y})^2}} \quad (-1 \leq r_{xy} \leq 1)$$

これの分子を偏差積と呼び、共分散と呼ぶ

x_i を $z_i = (x_i - \bar{x})/S_x$, y_i を $w_i = (y_i - \bar{y})/S_y$ と標準化したデータ z_i, w_i を考えると、定義から

$$r_{xy} = \frac{1}{n}\sum z_i w_i$$

・$-1 \leq r_{xy} \leq 1$ の証明

$$\frac{1}{n}\sum_{i=1}^{n}(z_i \pm w_i)^2 = \frac{1}{n}\sum_{i=1}^{n}(z_i^2 \pm 2z_i w_i + w_i^2)$$

$$= \frac{1}{n}\sum_{i=1}^{n} z_i^2 \pm \frac{2}{n}\sum_{i=1}^{n} z_i w_i + \frac{1}{n}\sum_{i=1}^{n} w_i^2$$

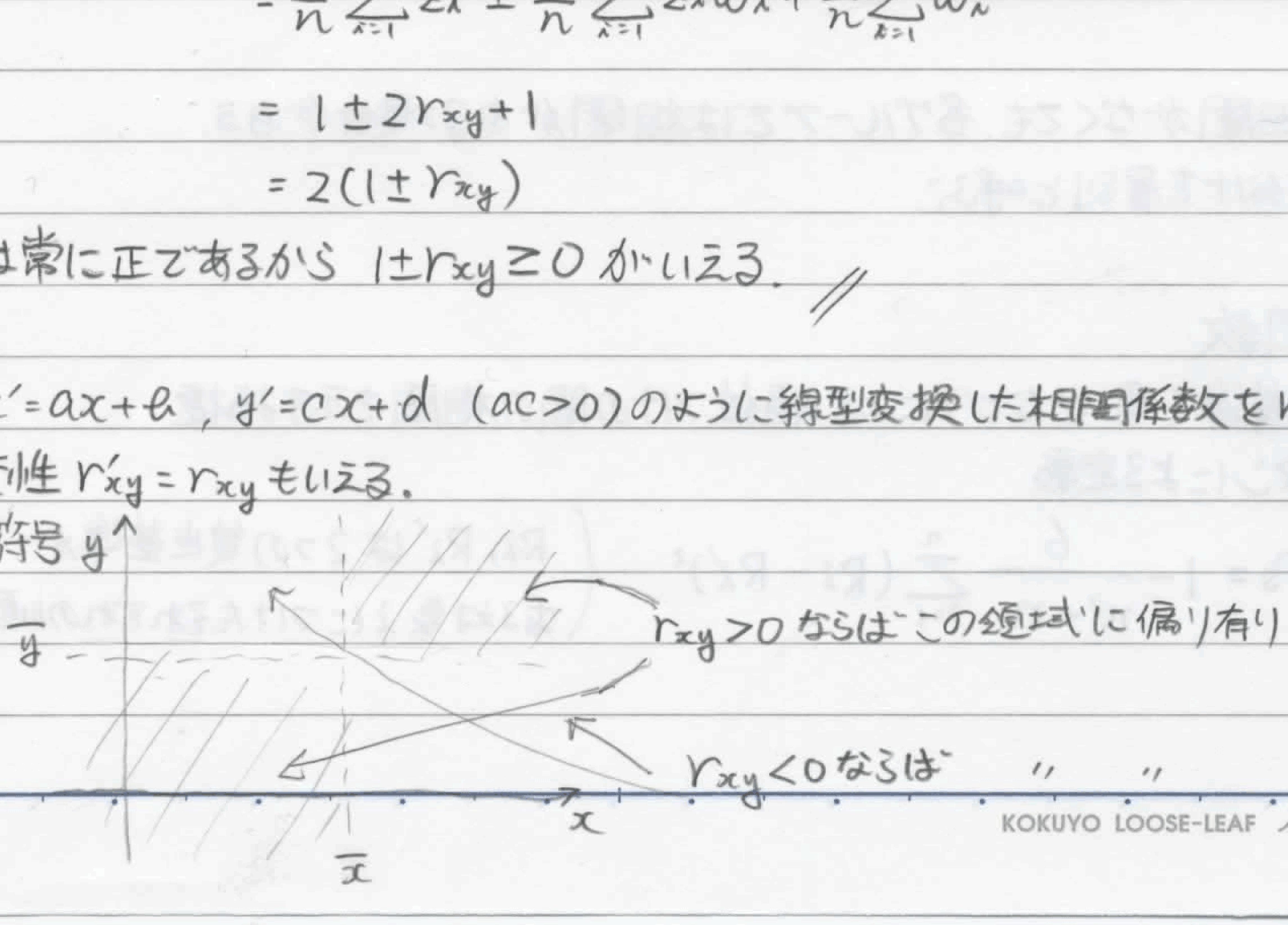
$$= 1 \pm 2r_{xy} + 1$$

$$= 2(1 \pm r_{xy})$$

左辺は常に正であるから $1 \pm r_{xy} \geq 0$ がいえる。//

更に $x' = ax + b$, $y' = cx + d$ $(ac > 0)$ のように線型変換した相関係数を r'_{xy} とすると線型不変性 $r'_{xy} = r_{xy}$ もいえる。

分子の符号



影の部分やそうでない部分に偏りが見られない場合 r_{xy} は0に近づく.

$r_{xy}=\pm 1$ をとるのは

$$y_i=\pm\frac{S_y}{S_x}(x_i-\bar{x})+\bar{y}$$

この時, 全てのデータが

$$y=\pm\frac{S_y}{S_x}x+\bar{y}\mp\frac{S_y}{S_x}\bar{x}$$

で示される直線上にある.

$r=1$ の時

$\frac{S_y}{S_x}>0$ であり, 正の完全相関という

$r=-1$ の時

$\frac{S_y}{S_x}<0$ であり, 負の完全相関という.

・「相関関係がある」≠「因果関係がある」

例 $y=(x-8)^2$

・みかけ上の相関

散布図を見ると, x, yに相関関係があるように見えるが, それはx, zの相関と, y, zの相関によるものであるというもの.

・偏相関係数

x, y, zの変数がある場合, zの影響を除いた下での, x, yの相関係数

$$r_{xy\cdot z}=\frac{r_{xy}-r_{xz}r_{yz}}{\sqrt{1-r_{xz}^2}\sqrt{1-r_{yz}^2}}$$

・層別

全体では相関がなくても, 各グループでは相関がある場合がある.

そのグループ分けを層別と呼ぶ.

・順位相関関数

二つの質的基準(男・女など)による順位づけの間の相関を示す指標

・スピアマンによる定義

$$r_s=1-\frac{6}{n^3-n}\sum_{i=1}^{n}(R_i-R_i')^2 \quad \left(\begin{array}{l}R_i, R_i' \text{は2つの質的基準が}\\ \text{ある対象}i\text{につけたそれぞれの順位}\end{array}\right)$$

Date . . No.

。ケンドールによる定義

正順，すなわち $R_i>R_j$, $R_i'>R_j'$，あるいは $R_i<R_j$, $R_i'<R_j'$ の場合は $+1$

逆順 〃 $R_i>R_j$, $R_i'<R_j'$ 〃 $R_i<R_j$, $R_i'>R_j'$ 〃 -1

$+1$を与えた対の数をG，-1を与えた対の数をHとする．

$$r_k=\frac{G-H}{n(n-1)/2}$$ ← 対の全個数

・自己相関係数

$x_1 \cdots x_n$を時系列とする．

遅れhの自己相関関数は

$$r_h=\frac{\sum_{i=1}^{n-h}(x_i-\bar{x})(x_{i+h}-\bar{x})/(n-h)}{\sum_{i=1}^{n}(x_i-\bar{x})^2/n} \quad \left(\bar{x}=\frac{x_1+\cdots+x_n}{n}\right)$$

。直線のあてはめ

2変数x, yにおいて，一方xが他方yを決定する関係があるとき，xを独立変数，yを従属変数と呼ぶ．

この場合2章にみられた回帰という見方や方法でx, yの関係を扱う．

x, yの関係が直線の場合，一般に

$$y=bx+a$$ … 回帰方程式，回帰直線　b：偏回帰係数

と表すことができる．しかし，b, aが客観的に定まらない

そこで使うのが

。最小二乗法

$$L=\sum_{i=1}^{n}\{y_i-(bx_i+a)\}^2$$

$y-\bar{y}=b(x-\bar{x})$とも表わせるから

点$(\bar{x},\bar{y})$を通る傾きbの直線

を最小にするa, bを求める．

これに偏微分などを用いると

$$\begin{cases} na+(\sum x_i)b=\sum y_i \\ (\sum x_i)a+(\sum x_i^2)b=\sum x_iy_i \end{cases}$$ … 正規方程式

$$\therefore \begin{cases} a=\bar{y}-b\bar{x} \\ b=\dfrac{\sum x_iy_i-n\bar{x}\bar{y}}{\sum x_i^2-n\bar{x}^2}=\dfrac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} \end{cases}$$

$$\left(\because \begin{array}{l}\sum(x_i-\bar{x})(y_i-\bar{y})=\sum x_iy_i-n\bar{x}\bar{y} \\ \sum(x_i-\bar{x})^2=\sum x_i^2-n\bar{x}^2\end{array}\right)$$

相関係数r_{xy}を考えると

$$b=r_{xy}\frac{S_y}{S_x}$$

となる．

No. Date

得られた $y = bx + a$ に x_i を代入したyの値を $\hat{y}_i$ とし、実際に得られている y_i との差を

$$d_i = y_i - \hat{y}$$

とする。すると

$$\sum d_i^2 = \sum (y_i - \hat{y})^2 = (1 - r^2) \sum (y_i - \bar{y})^2 \quad \cdots \text{(a)}$$

が成り立つ。

この式からわかるように r^2 が1に近いほど x が y を決定する強度が強くなる。

この r^2 を決定係数と呼ぶ。

(a)の証明

回帰係数 b の定義より

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = \sum b(x_i - \bar{x})^2$$

$$\therefore \sum (x_i - \bar{x})\{(y_i - \bar{y}) - b(x_i - \bar{x})\} = 0$$

$$\therefore \sum (x_i - \bar{x})\{y_i - (bx_i + a)\} = 0 \qquad (\because a = \bar{y} - b\bar{x})$$

$$\therefore \sum (\hat{y}_i - \bar{y})(y_i - \hat{y}) = 0 \qquad (\because \hat{y}_i - \bar{y} = b(x_i - \bar{x}))$$

となる。$y_i - \bar{y} = (y_i - \hat{y}) + (\hat{y}_i - \bar{y})$ として、この結果を使うと、

$$\sum (y_i - \bar{y})^2 = \sum (y_i - \hat{y})^2 + \sum (\hat{y}_i - \bar{y})^2 \quad \cdots (*)$$

回帰直線より $\hat{y}_i - \bar{y} = (rS_y/S_x)(x_i - \bar{x})$ であるから、

$$\sum (\hat{y}_i - \bar{y})^2 = (r^2 \cdot nS_y^2 / nS_x^2) \sum (x_i - \bar{x})^2$$
$$= r^2 \cdot nS_y^2 = r^2 \cdot \sum (y_i - \bar{y})^2$$

これを(*)に代入して(a)を得る //

・平面のあてはめ

yを決定する変数を2つ以上とする。重回帰。

第4章

・標本空間と事象

さいころを1回投げた場合、可能な結果は、1,2,3,4,5,6のみ.

このような可能な結果を標本点 ω、その全体の集合を標本空間 Ω と呼ぶ.

標本空間が有限の r 個の標本点からなる場合、事象は 2^r 個存在する.

根元事象 … ただ一つの標本点からなる分解できない事象　さいころ $\{1\},\{2\}$

複合事象 … 複数の標本点を含み、2つ以上の根元事象に分解可能なもの　$\{1,2\}$

・ラプラスの定義が使えない時、つまり各標本点が「同程度に確らしく」起りやすいと考えられない場合には、頻度説を用いる.

一般に事象Aを生み得る実験を n 回繰り返してAが n_A 回出るとすると、$n\to\infty$ のとき

$$n_A/n \to \alpha$$

となるならば、$P(A)=\alpha$ と定義する.

・確率の公理主義的定義

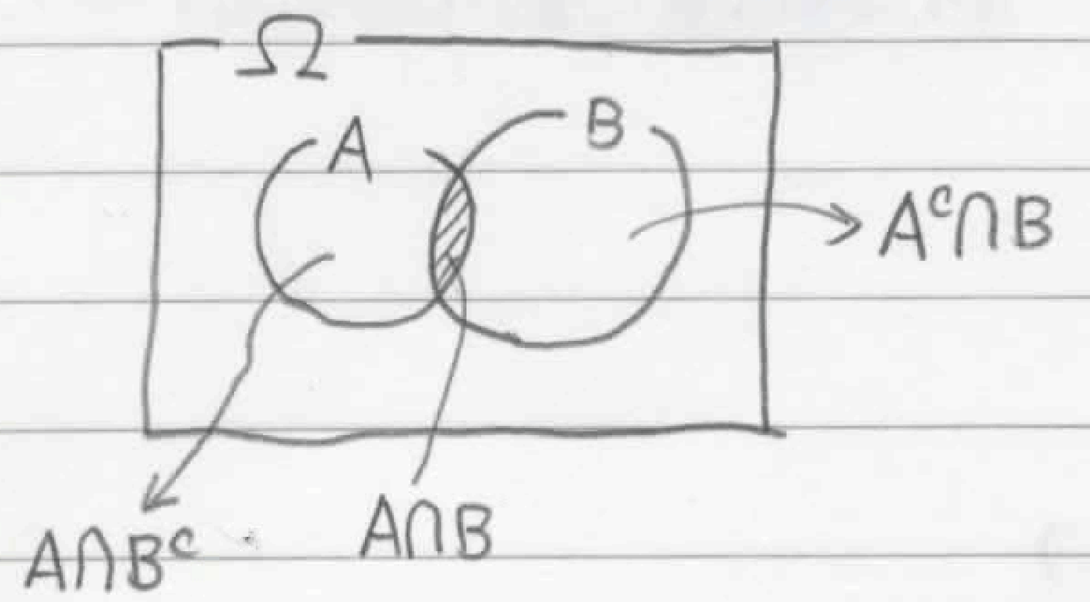
公理

(a) 全ての事象Aに対して $0\leq P(A)\leq 1$

(b) $P(\Omega)=1$

(c) 互いに排反な事象 A_1,A_2,A_3 … に対して

$$P(A_1\cup A_2\cup A_3\cup\cdots)=P(A_1)+P(A_2)+P(A_3)+\cdots$$



・加法定理

$A\cap B=\phi \rightarrow P(A\cup B)=P(A)+P(B)$ を加法定理と呼ぶ

$A\cap B\neq\phi$ の時

$$\begin{cases} P(A\cup B)=P(A\cap B^c)+P(A\cap B)+P(A^c\cap B) \\ P(A)=P(A\cap B^c)+P(A\cap B) \\ P(B)=P(A^c\cap B)+P(A\cap B) \end{cases}$$

$\therefore P(A\cup B)=P(A)+P(B)-P(A\cap B)$

加法定理

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(A\cap B)-P(B\cap C)-P(C\cap A)+P(A\cap B\cap C)$$

・条件付確率

Bを条件とするAの条件付確率　$P(A|B)$と表す。

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

例　$P(1|白)=\dfrac{P(白かつ1)}{P(白)}$

$P(A\cap B)=P(B)\cdot P(A|B)$ … 乗法定理

… (*)

・独立性

$P(A)=P(A|B)$ である時、独立であるという。

(*)に代入すると

$P(A\cap B)=P(A)\cdot P(B)$　　　($P(A\cap B\cap C)=P(A)\cdot P(B)\cdot P(C)$)

・ベイズの定理

結果Aの原因をH_iとする。

ベイズの定理は$P(H_i|A)$を計算するためのもの。

$H_1, H_2, \cdots, H_k$ は互いに排反

$H_1\cup H_2\cup\cdots\cup H_k=\Omega$ であったとする。

この時

$$P(H_i|A)=\frac{P(H_i)\cdot P(A|H_i)}{\sum P(H_j)\cdot P(A|H_j)}$$

$P(A)$が簡単に出せないのは
H_iを得た後にAが得られる。
H_iを選ぶこととAを選ぶことは
同じではない
例．教科書に書いてある例がよい

$P(H_i)$は事前確率，$P(H_i|A)$は事後確率と呼ばれる。

証明

$$P(H_i|A)=\frac{P(H_i\cap A)}{P(A)}=\frac{P(A|H_i)\cdot P(H_i)}{P(A)}$$ (∵乗法定理など)

… (#)

集合の分配法則より

$A=A\cap\Omega$

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

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

$A\cap H_1, A\cap H_2, \cdots, A\cap H_k$ は互いに排反であるから，

$P(A)=\sum P(A\cap H_j)=\sum P(H_j)\cdot P(A|H_j)$ … 全確率の定理

これを(#)に代入すれば上記の式が得られる。

ベイズの定理を基礎とする統計的方法をベイズ統計学という。

番号

可算集合 … 自然数全体と一対一の対応がつけられる、すなわち順番のつけられる集合.

Date . . No.

第5章

・あるものがとる各値に対しそれぞれ確率が与えられている確率変数という.
確立変数は、Xのように大文字を用いて表す.

・可算集合$\{x_1, x_2, \cdots\}$の中の値を取る確率変数Xは離散型と呼ばれる.
それぞれの値の確率は

$$P(X=x_k)=f(x_k)$$

で表され、これをXの確率分布という. fは教科書のサイコロの例のように.

$$f(x_k) \geq 0 \quad (k=1,2\cdots), \quad \sum_{k=1}^{\infty} f(x_k)=1$$

を満たす.
このfを離散型の確率分布という.

・連続型の確率変数
確率変数Xのとる値が関数$f(x)$によって

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

と表される場合, Xは連続型の確率分布をもつという.

これは

$$\text{全ての}x\text{に対し } f(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x)dx = 1$$

を満たす. この$f(x)$をXの確率密度関数と呼ぶ.

注 $a=b$とすれば

$$P(X=a)=0$$

であるから、一点の確率は0となる. これが離散型との大きな違い.

・指数分布
$\lambda > 0$として

$$\begin{cases} f(x) = \lambda e^{-x} & (x \geq 0) \\ f(x) = 0 & (x < 0) \end{cases}$$

とすると、$f(x) \geq 0$で

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} \lambda e^{-\lambda x}dx = 1$$

と規格化されているから, 確率分布の性質を満たす. 定積分をしてみると、

$$P(a \leq X \leq b) = \int_a^b \lambda e^{-\lambda x}dx = e^{-\lambda a} - e^{-\lambda b}$$

ということがわかる.
これが使われるのは ある災害が起ってから次の災害が起るまでの時間X, 電球が偶発的に切れるまでの寿命Xなど待ち時間を確率変数とした時.
ある災害が起った時点を$x=0$として計算する(?)

No. Date

・一様分布（連続型）

これの密度関数は　　　　　　　　　　　　ルーレットエリア指定

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

一様乱数やルーレットはこれに従う。

・累積分布関数

確率変数Xに対して，xを実数とする時 x以下の確率

$$F(x)=P(X \leqq x)$$

をXの累積分布関数と呼ぶ。

{例えば飛行機の積載重量Xと安全確率，（最大積載量をxとした）}

・連続型の場合は

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

と表される。

$F(x)$は確率分布の他の表し方であり，$F(x)$でも$f(x)$でも確率分布を表すことができる。

先程の指数分布の累積分布関数は

$$\begin{cases} F(x)=\int_{-\infty}^{x} f(u)d(u)=\int_{-\infty}^{x} 0\cdot du=0 & (x<0) \\ F(x)=\int_{-\infty}^{x} f(u)d(u)=\int_{0}^{x} f(u)du=1-e^{-\lambda x} & (x \geqq 0) \end{cases}$$

となる。

（時点tにおいて災害がおこっていない確率は
$P(X>t)=1-F(t)=e^{-\lambda t}$）

[一様分布は

$$\begin{cases} F(x)=0, & x<0 \\ F(x)=x, & 0 \leqq x \leqq 1 \\ F(x)=1, & x>1 \end{cases}$$
]

・離散型の場合は

$$F(x)=\sum_{u \leqq x} f(u)$$

たとえばXが整数値のみをとるとし，その離散的な確率分布を$P(x=k)=f(k)$とすると，εがきわめて小さければ（$\varepsilon>0$），定義から

$$F(k)-F(k-\varepsilon)=P(k-\varepsilon<X \leqq k)=P(X=k)=f(k)$$

となる。このように，$F(x)$は整数点kで不連続となり，その点におけるジャンプの高さ（増加分）が$f(k)$となる。ただし，右からは連続。

・累積分布関数の性質

◦ 広義単調増加　$x_1<x_2$ ならば $F(x_1) \leqq F(x_2)$

◦ 範囲　$x \to \infty$　$F(x) \to 1$

　　　　$x \to -\infty$　$F(x) \to 0$

◦ 右連続　$\varepsilon \to 0+0$　$F(x+\varepsilon) \to F(x)$

Xを1個のさいころを振ったときあらわれる目とすると

$E(X)=\frac{7}{2}$

◦ モードとメディアン

確率分布を代表する値

モード… $f(x)$ を最大にする x

メディアン… $P(X \leq x)=\frac{1}{2}$ となるx

◦期待値と分散

・期待値 $E(X)$

離散型

$$E(X)=\sum_{x} x f(x)$$

連続型

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

期待値の演算Eについて、重要な性質

(a) $E(c)=c$

(b) $E(X+c)=E(X)+c$

(c) $E(cX)=cE(X)$

(d) $E(X+Y)=E(X)+E(Y)$

教科書P96に書いてある例えが非常にわかりやすい。

◦ 分散と標準偏差

期待値が同じでも、異なった確率分布はいくらでもある。確率のばらつきや集中度を表すのが分散と標準偏差である。

$E(X)=\mu$とする。

$$V(X)=E\{(X-\mu)^2\}$$

と定義し、これをXの分散と呼ぶ。

・$V(X)\geq 0$であり、$V(X)$の値が大きいほどXのばらつきは大きい。

分散の計算は定義から

離散型

$$V(X)=\sum_{x}(x-\mu)^2 f(x)$$

連続型

$$V(X)=\int_{-\infty}^{\infty}(x-\mu)^2 f(x) dx$$

しかし、

$$\begin{aligned} E\{(X-\mu)^2\} &= E(X^2-2\mu X+\mu^2) \\ &= E(X^2)-2\mu E(X)+\mu^2 && (\because (a)\sim(d)) \\ &= E(X^2)-\mu^2 && (\because E(X)=\mu) \\ &= E(X^2)-\{E(X)\}^2 && (\because E(X)=\mu) \end{aligned}$$

を用いた方が楽である。

$D(X)=\sqrt{V(X)}$

と定義し、これをXの標準偏差と呼ぶ

分散の演算Vについて、重要な性質

(a) $V(c)=0$

(b) $V(X+c)=V(X)$

(c) $V(cX)=c^2V(X)$

・標準化変数

$$Z=\frac{\{X-E(X)\}}{\sqrt{V(X)}}$$

と定義する。Zを標準化変数と呼ぶ。

$E(Z)=0 \quad V(Z)=1$

となる。

・モーメントとモーメント母関数

・歪度・尖度

確率分布の様子を表すもの

・歪度

右の裾が長く、歪んでいるといった非対称性の方向、およびその程度を表す。

$$\alpha_3=\frac{E\{(X-\mu)^3\}}{\sigma^3} \quad \begin{pmatrix} E(X)=\mu \\ D(X)=\sigma \end{pmatrix}$$

と定義し、α_3を歪度と呼ぶ

$$\begin{cases} \alpha_3>0 & \text{右の裾が長い} \\ \alpha_3<0 & \text{左の裾が長い} \\ |\alpha_3| & \text{非対称性の程度} \end{cases}$$

しかし、先ほどと同様で、

$$E\{(X-\mu)^3\}=E(X^3)-3\mu E(X^2)+2\mu^3$$

を用いることができる。

・尖度

中心の周囲の部分の尖り具合を表す。

$$\alpha_4=\frac{E\{(X-\mu)^4\}}{\sigma^4} \quad \begin{pmatrix} E(X)=\mu \\ D(X)=\sigma \end{pmatrix}$$

と定義し、α_4-3を尖度と呼ぶ。

$$\begin{cases} \alpha_4 - 3 > 0 & \text{正規分布より尖っている} \\ \alpha_4 - 3 < 0 & \text{正規分布より丸く鈍い形} \end{cases}$$

先ほどと同様

$$E\{(X-\mu)^4\} = E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 4\mu^3 E(X) + \mu^4$$
$$= E(X^4) - 4\mu E(X^3) + 6\mu^2 E(X^2) - 3\mu^4$$

を用いることができる。

・モーメント

◦ $\mu_r = E(X^r)$ $(\mu_1 = E(X))$

をXの(原点まわりの)r次のモーメント、又は積率と呼ぶ

◦ $\mu_r' = E\{(X-\mu)^r\}$ $(\mu = E(X))$ $(\mu_2' = V(X))$

をXの期待値(平均)のまわりのr次のモーメントと呼ぶ

◦ $\alpha_r = E\left\{\frac{(X-\mu)}{\sigma}\right\}^r$ $(\sigma = D(X))$ $\begin{pmatrix} r=3 & \text{歪度} \\ r=4 & \text{尖度} \end{pmatrix}$

をXのr次の標準化モーメントと呼ぶ.

全ての次数のモーメントを指定すれば、それによって一つの確率分布が決定される。
すべての次数のモーメントを生成するモーメント母関数を

$$M_X(t) = E(e^{tX})$$

と定義する

$$\begin{cases} \text{離散型} \\ \quad M_X(t) = \sum_x e^{tx} f(x) \\ \text{連続型} \\ \quad M_X(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\,dx \end{cases}$$

モーメント母関数について

$$M_X^{(r)}(0) = \mu_r$$

が成り立つ.

証明

e^xの展開式 $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$ に $x = tX$ を代入し、

$$e^{tX} = 1 + tX + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots$$

これの両辺の期待値をとれば

$$M_X(t) = 1 + tE(X) + t^2E(X^2)/2! + t^3E(X^3)/3! + \cdots$$
$$= 1 + \mu_1 t + (\mu_2/2!)t^2 + (\mu_3/3!)t^3 + \cdots$$

となり、明らか.

No. Date

第6章

○離散型の確率分布

○二項分布, ベルヌーイ分布

2種類の可能な結果(SとF)を生じる実験あるいは観測があり, それらの確率をそれぞれ p, $1-p$ とする.
これを, 同じ条件でかつ独立に n 回繰り返すことを考える.
これを, ベルヌーイ試行と呼ぶ。
いま, Sが x 回, Fが $n-x$ 回生じるとすると, その確率は

$$f(x) = {}_nC_x p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \cdots, n$$

である. この確率分布を二項分布と呼び, $Bi(n, p)$ と表す.
特に $Bi(1, p)$ をベルヌーイ分布と呼ぶ.
これが確率分布であることは

$$\sum_x {}_nC_x p^x (1-p)^{n-x} = \{p + (1-p)\}^n = 1$$

であることから確かめられる.
X が二項分布 $Bi(n, p)$ に従っているならば

$$E(X) = np \quad V(X) = np(1-p)$$

証明

$$\begin{aligned}
E(X) &= \sum_x x \cdot {}_nC_x p^x q^{n-x} \\
&= \sum_x n \cdot {}_{n-1}C_{x-1} p \cdot p^{x-1} q^{n-x} \\
&= np \sum_y {}_{n-1}C_y p^y q^{(n-1)-y} \quad (y = x-1) \\
&= np
\end{aligned}$$

//

○ポアソン分布

n が大きくなり, p が小さくなると ${}_nC_x p^x (1-p)^{n-x}$ の計算が困難である.
このような時, 以下のような定理が使える
$np \to \lambda$ である時, $n \to \infty$, $p \to 0$ となる極限では, 各 x について

$${}_nC_x p^x (1-p)^{n-x} \to \frac{e^{-\lambda}\lambda^x}{x!}$$

が成り立つ.
これをポアソンの小数の法則という.

$$\lambda > 0, \quad f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad (x = 0, 1, 2, \cdots)$$

この確率分布をポアソン分布と呼び, $P_o(\lambda)$ で表す.
これが確率分布であることは

$$\sum_x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \underbrace{\sum_x \frac{\lambda^x}{x!}}_{e^{\lambda}\text{の展開式}} = e^{-\lambda} \cdot e^{\lambda} = 1$$

であることから確かめられる.
X がポアソン分布 $P_o(\lambda)$ に従うならば

$$E(X) = \lambda \quad V(X) = \lambda$$

証明

$$\begin{aligned}
E(X) &= \sum_x x \cdot \frac{e^{-\lambda}\lambda^x}{x!} = \lambda e^{-\lambda} \sum_x \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \sum_y \frac{\lambda^y}{y!} \\
&= \lambda e^{-\lambda} e^{\lambda} \\
&= \lambda
\end{aligned}$$

//

Date . . No.

- 一様分布

$f(x) = 1/N$, $x = 1, 2, \cdots, N$ （Nは正整数）

を，1, 2 … N上の離散一様分布という。

- 連続型の確率分布
  - 指数分布

$$\begin{cases} f(x) = \lambda e^{-\lambda x} & x \geq 0 \\ f(x) = 0 & x < 0 \end{cases}$$

で定義される確率分布を指数分布と呼ぶ

Xがこの確率分布に従うならば，

$$E(X) = \frac{1}{\lambda}, \quad V(X) = \frac{1}{\lambda^2}$$

KOKUYO LOOSE-LEAF ノ-836B 6mm ruled×36 lines

第7章

◦ 同時確率分布

確率変数 X, Y があり, $X=x$ であり, $Y=y$ である確率は

$$P(X=x, Y=y)=f(x,y) \quad \text{(離散型)}$$

と表され これを同時確率分布と呼ぶ

$f(x,y)$ は

離散型の場合は

$$f(x,y)\geq 0, \ \sum_x\sum_y f(x,y)=1$$

連続型の場合は

$$f(x,y)\geq 0, \ \iint_S f(x,y)\,dx\,dy=1$$

この時 $f(x,y)$ は同時確率密度関数と呼ぶ

(S ← 標本空間全積)

を満たす.

又,

離散型の場合は

$$P((X,Y)\in A)=\sum_A\sum f(x,y)$$

連続型の場合は

$$P(a\leq X\leq b, c\leq Y\leq d)=\int_c^d\int_a^b f(x,y)\,dx\,dy$$

となる.

S: ユークリッド空間

事象 A: S の部分集合

◦ 周辺確率分布

$$g(x)=\sum_y f(x,y), \quad h(y)=\sum_x f(x,y) \quad \text{(離散型)}$$

と定義し, それぞれを X, Y の周辺確率分布と呼ぶ。

$$g(x)=\int_{-\infty}^{\infty} f(x,y)\,dy, \quad h(y)=\int_{-\infty}^{\infty} f(x,y)\,dx \quad \text{(連続型)}$$

と定義し, これらを周辺確率密度関数と呼ぶ.

※同時確率分布 $\rightleftarrows$ 周辺確率分布

※ $E(X+Y)=E(X)+E(Y)$ の証明

$$\begin{aligned}E(X+Y)&=\iint (x+y)f(x,y)\,dx\,dy\\&=\iint x f(x,y)\,dx\,dy+\iint y f(x,y)\,dx\,dy\\&=E(X)+E(Y) \quad /\!/\end{aligned}$$

◦ 共分散

$$V(X+Y)\neq V(X)+V(Y)$$

じゃあ $V(X+Y)$ は?

これ要チェックだと思った.

$$\begin{aligned}V(X+Y)&=E\left[\{(X+Y)-E(X+Y)\}^2\right]\\&=E\left[\{(X+Y)-(E(X)+E(Y))\}^2\right]\\&=E\left[\{(X-\mu_X)+(Y-\mu_Y)\}^2\right] \quad (E(X)=\mu_X,\ E(Y)=\mu_Y \text{ とおいた})\\&=E\left[(X-\mu_X)^2+(Y-\mu_Y)^2+2(X-\mu_X)(Y-\mu_Y)\right]\\&=E\{(X-\mu_X)^2\}+E\{(Y-\mu_Y)^2\}+2E\{(X-\mu_X)(Y-\mu_Y)\}\\&=V(X)+V(Y)+2\mathrm{Cov}(X,Y)\end{aligned}$$

Date . . . No.

$Cov(X, Y) = E\{(X-\mu_X)(Y-\mu_Y)\}$

と定義し、これをX、Yの共分散と呼ぶ。

$\begin{cases} Cov(X,Y) > 0 & X、Yは大小が同傾向 \\ Cov(X,Y) < 0 & 〃 \quad 逆傾向 \end{cases}$

$$
\begin{aligned}
Cov(X,Y) &= E(XY - \mu_X Y - \mu_Y X + \mu_X\mu_Y) \\
&= E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X\mu_Y \\
&= E(XY) - \mu_X\mu_Y \\
&= E(XY) - E(X)E(Y)
\end{aligned}
$$

$\because (*)$

◦相関係数

共分散 $Cov(X, Y)$ が、X、Yの関係の方向を表すのに対し、その強さの程度を判断する基準を

$$\rho_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}, \quad -1 \leq \rho_{XY} \leq 1$$

と定義し、これを相関係数と呼ぶ。

$\begin{cases} \rho_{XY} > 0 & X、Yは大小が同傾向 \\ \rho_{XY} < 0 & X、Yは大小が逆傾向 \\ |\rho| & この値が大きくなるとX、Yの関係が確定的なものになっていく。 \end{cases}$

$\rho = \pm 1$ の時、X、Yの間には厳密な1次式の関係が成り立つ。

$Y = aX + b$ （$\rho = 1$ ならば $a > 0$，$\rho = -1$ ならば $a < 0$）

$\rho = 0$ の時、X、Yは無相関であるという。

$\begin{cases} 離散型では \\ Cov(X,Y) = \sum_x \sum_y (x-\mu_X)(y-\mu_Y)\cdot f(x,y) \\ 連続型では \\ Cov(X,Y) = \iint_S (x-\mu_X)(y-\mu_Y)\cdot f(x,y)\,dx\,dy \end{cases}$

と表されるが、$(*)$ を用いて計算するのが通常である。

$$\left(\begin{aligned} &E(XY) = \sum_x \sum_y xy\cdot f(x,y) \quad (離散型) \\ &E(X,Y) = \iint_S xyf(x,y)\,dx\,dy \quad (連続型) \end{aligned}\right)$$

◦条件付確率

$$\begin{cases} P(X=x \mid Y=y) = \dfrac{P(X=x, Y=y)}{P(Y=y)} \\ P(Y=y \mid X=x) = \dfrac{P(X=x, Y=y)}{P(X=x)} \end{cases}$$

と定義し、Y(X)がy(x)と与えられたときのX(Y)の条件付確率密度関数と呼ぶ。

確率分布の条件について

$g(x|y) = f(x,y)/h(x) \quad h(y|x) = f(x,y)/g(x)$

と定義すると、

No. Date

$\sum_x g(x|y) = \sum_x f(x,y)/h(y) = h(y)/h(y) = 1$

となり確かに満たしている。

。条件付期待値

$$\begin{cases} E(X|y) = \sum_x x \cdot g(x|y) = \mu_{X|y} \\ E(Y|x) = \sum_y y \cdot h(y|x) = \mu_{Y|x} \end{cases}$$

と定義される。

連続型ならば
$\sum$ を $\int$ に変える

。条件付分散

$$\begin{cases} V(X|y) = \sum_x (x-\mu_{X|y})^2 g(x|y) \\ V(Y|x) = \sum_y (y-\mu_{Y|x})^2 h(y|x) \end{cases}$$

と定義される。

。独立な確率変数

同時確率分布においてあらゆる x, y について、条件

$$f(x,y) = g(x)h(y)$$

が成り立つならば X、Y は互いに独立であるという。

この時、

$$g(x|y) \equiv g(x), \quad h(y|x) \equiv h(y)$$

が成り立つ。

又、X、Y が独立ならば

$$E(XY) = E(X)E(Y)$$

証明

$$E(XY) = \sum_x \sum_y xy \cdot f(x,y) = \sum_x \sum_y xy \cdot g(x)h(x)$$
$$= E(X)E(Y)$$

が成り立ち、これを用いると

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0$$

となり、相関係数 $\rho = 0$ であることがわかる。

注　無相関 $\not\Leftarrow$ 独立

。分散の加法性

X、Y が独立である時

$$V(X \pm Y) = V(X) + V(Y)$$

が成り立つ

証明

$$V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$$
$$= V(X) + V(Y) + 2\rho_{XY} D(X) D(Y)$$

X、Y が独立の時 $\rho_{XY} = 0$

$X_1, \cdots, X_n$ が独立のときには $V(X_1 + \cdots + X_n) = V(X_1) + \cdots + V(X_n)$ が成り立つ。

Date . . No.

・同一分布

$E_1(X_1)=E(X_2)=\cdots=E(X_n)=\mu$

$V(X_1)=V(X_2)=\cdots=V(X_n)=\sigma^2$

である時

$E(X_1+X_2+\cdots+X_n)=n\mu$

$V(X_1+X_2+\cdots+X_n)=n\sigma^2$

従って、標準偏差

$D(X_1+X_2+\cdots+X_n)=\sqrt{n}\sigma$

・相加平均

$$\bar{X}=\frac{X_1+X_2+\cdots+X_n}{n}$$

とすると

$E(\bar{X})=\mu$

$V(\bar{X})=\dfrac{\sigma^2}{n}$ $(\because V(cX)=c^2V(X))$

・和の確率分布 たたみこみ

独立な確率変数 X, Y の確率分布(密度関数)を $g(x)$, $h(y)$ とした時.

$X+Y$ の確率分布(密度関数)は

離散型

$$k(z)=\sum_x g(x)h(z-x)$$

連続型

$$k(z)=\int_{-\infty}^{\infty} g(x)h(z-x)dx$$

と表される.

関数 g, h から k を作る数学操作をたたみこみと呼び $k=g*h$ と書く

・二項分布

$Bi(n,p)*Bi(m,p)=Bi(n+m*p)$

・ポアソン分布

$Po(\lambda)*Po(\mu)=Po(\lambda+\mu)$

たたみこみの結果、同種の確率分布が得られる時、その確率分布族は再生的であるという.

No. Date

第8章

○大数の法則

ベルヌーイ試行において、試行回数が多ければ多いほど、観測される成功率が真の成功率に近づいていくというもの。

$E(x) = np$

$E(\frac{x}{n}) = p$ … 真の成功率

○中心極限定理

母集団分布が何であっても、和 $X_1 + X_2 + \cdots + X_n$ の確率分布の形は、n が大なるときは、大略正規分布と考えてよいというもの。

第9章

読んで下さい

第10章

◦正規分布の性質 (ガウス分布)

◦ 確率変数Xが正規分布に従うとき，その密度関数は

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad \cdots \quad x=\mu に関して対称$$

と表し、$N(\mu,\sigma^2)$と表す。又、

$E(X)=\mu$, $V(X)=\sigma^2$

である。 ← メディアンとモードと一致する。

$$\begin{cases} aX+b は N(a\mu+b, a^2\sigma^2) \\ X,Y が N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2) に従うとき. X+Y は N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2) に従う. \\ \qquad X-Y は N(\mu_1-\mu_2, \sigma_1^2+\sigma_2^2) に従う. \end{cases}$$

※ $\mu=0$, $\sigma^2=1$の正規分布を特に標準正規分布と呼ぶ
この時. 密度関数を$\phi(x)$, 累積分布関数を$\Phi(x)$で表す.

※ $X_1, X_2 \cdots X_n$はそれぞれ独立かつ$N(\mu,\sigma^2)$に従うとき、統計量

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$

は標準正規分布$N(0,1)$に従う.

• 各X_iが正規分布$N(\mu,\sigma^2)$に従うならば，$\bar{X}$は正規分布$N(\mu,\sigma^2/n)$に従う.
$\bar{X}$は単独のX_iよりも分散が小さく. 測定の正確さがあがる.

• 標本分散は

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2 \qquad (n\to\infty の時 \ s^2 \fallingdotseq \sigma^2)$$

と定義され、母分散をσ^2とすると

$E(s^2)=\sigma^2$

を満たす.

No. Date

◦ χ^2分布

$Z_1, Z_2, \cdots, Z_k$を独立な, 標準正規分布$N(0,1)$に従う確率変数とする.

$$\chi^2 = Z_1^2 + Z_2^2 + \cdots + Z_k^2$$

とし、確率変数χ^2が従う確率分布を自由度kのχ^2分布という.

自由度kのχ^2分布を$\chi^2(k)$で表す.

正規母集団からの標本$X_1, X_2, X_3 \cdots X_n$の標本分散をS^2とすると

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$

であり, これは自由度$(n-1)$のχ^2分布$\chi^2(n-1)$に従う.

・t分布

$$Z = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}}$$

はσ^2がなければ計算できないので,

t統計量

$$t = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} \quad \cdots (*)$$

分母を標本平均の標準誤差という.

を定義する. これが従う確率分布をt分布と呼ぶ. $(*)$は自由度$n-1$のt分布$t(n-1)$に従う.

・t分布の定義

(a) Zは標準正規$N(0,1)$に従う.

(b) Yは自由度kのχ^2分布$\chi^2(k)$に従う.

(c) ZとYは独立である.

という(a)〜(c)の条件を満たす確率変数Y, Zを用いて,

確率変数tを

$$t = \frac{Z}{\sqrt{Y/k}}$$

と定義すると, tが従う確率分布を自由度kのt分布という.

$n \to \infty$の時 t分布は標準正規分布$N(0,1)$に一致する.

第11章

母集団の確率分布を決めている定数 すなわち母数を標本 $X_1, \cdots, X_n$ から定めることを母数の推定という. 標本平均や標本分散のように母数を推定するために標本から求めた統計量を一般に推定量と呼ぶ. 推定量は母数の記号に ^ をつけて表す.
推定量のとりうる値の一つが実現したものを推定値と呼ぶ.

- 点推定

  母数 θ を推定する場合, それをある1つの $\hat{\theta}$ で指定する方法

  - 点推定の手順

    - モーメント法

      モーメントを用いる.

      母集団が正規分布 $N(\mu, \sigma^2)$ のとき, $(\mu_r = E(X^r))$

      $$\sigma^2 = \mu_2 - \mu_1^2 \qquad (V(X) = E(X^2) - \{E(X)\}^2)$$

      より

      $$\mu_1 = \mu, \ \mu_2 = \sigma^2 + \mu^2$$

      である. 一方, 標本から, μ_1, μ_2 を

      $$\hat{\mu}_1 = \sum X_i / n, \quad \hat{\mu}_2 = \sum X_i^2 / n$$

      で推定することを考える. (母モーメント) = (標本モーメント) とすると

      $$\hat{\mu} = \sum X_i / n \quad \hat{\sigma}^2 = \sum X_i^2 / n - \bar{X}^2$$

      が得られる

      このようにモーメントの推定によって, 母数の推定を行う手続きをモーメント法という.

    - 最尤法

      母数空間 … 母数がとりうる値の集合を Θ で表す.

      1をとる確率が p, 0をとる確率が $1-p$ のベルヌーイ分布 $Bi(1, p)$ が母集団分布を考える.
      この時推定すべき母数は p である. $X_1 = 1, X_2 = 1, X_3 = 1, X_4 = 1, X_5 = 0$ の $n = 5$ の標本が得られたとする. この標本が得られる確率は

      $$L(p) = p^4(1-p) \cdots$$ (Θ での p のもっともらしさを表す関数, 尤度関数と呼ぶ)

      であり, 「現実の標本は確率最大のものが実現した」という最尤原理に基づいて, 最尤推定値 最尤推定量を決定する. この場合では最尤推定値は $\hat{p} = 0.8$ である.

      一般に, 大きさ n の標本において, 未知の母数を一般に θ とした場合, 尤度関数は $X_1, X_2, \cdots, X_n$ の同時確率分布を θ の関数とみなしたもので, X_i の確率分布を $f(x, \theta)$ とすると, 尤度関数は

      $$L(\theta) = f(x_1, \theta) \cdot f(x_2, \theta) \cdots f(x_n, \theta)$$

      となる. (複数の母数を持つとき
      $L(\theta_1, \theta_2, \cdots, \theta_n) = \prod_{i=1}^{n} f(x_i; \theta_1, \theta_2 \cdots \theta_n)$)

P218の例をみておく.

・点推定の基準

・不偏性　$E(\hat{\theta})=\theta$ となる性質

・一致性　$n\to\infty$ で $\hat{\theta}_n\to\theta$

・漸近正規性

・有効性

・区間推定

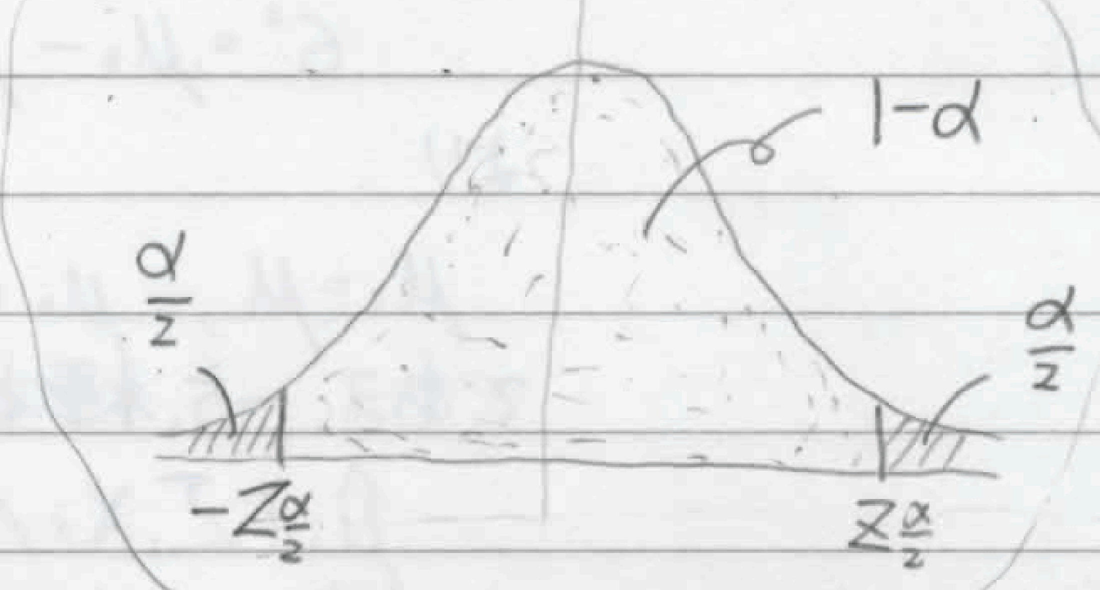
真のパラメーターの値が入る確率がある値 $1-\alpha$ 以上と保証される区間$[L, U]$をもとめるもの

$P(L\leqq\theta\leqq U)\geqq 1-\alpha$

下側信頼限界　信頼係数

上側信頼限界

何故 $\frac{\alpha}{2}$ なのか？



区間$[L, U]$を $100(1-\alpha)\%$ 信頼区間と呼ぶ

・母平均の信頼区間

μの信頼区間は

$$\left[\bar{X}-z_{\frac{\alpha}{2}}\cdot\sigma/\sqrt{n}\leqq\mu\leqq\bar{X}+z_{\frac{\alpha}{2}}\cdot\sigma/\sqrt{n}\right]$$

区間幅 $2z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$

パーセント点

$$\left[\bar{X}-t_{\frac{\alpha}{2}}(n-1)\ s/\sqrt{n}\leqq\mu\leqq\bar{X}-t_{\frac{\alpha}{2}}(n-1)\ s/\sqrt{n}\right]$$

$\frac{\sqrt{n}(\bar{X}-\mu)}{s}$ が自由度$(n-1)$のt分布$t(n-1)$に従うから

・母分散の信頼区間

σ^2の信頼区間は

$$\left[\frac{(n-1)s^2}{\chi^2_{\frac{\alpha}{2}}(n-1)},\ \frac{(n-1)s^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}\right]\leftarrow\frac{(n-1)s^2}{\sigma^2}$$

は自由度$n-1$のχ^2分布$\chi^2(n-1)$に従うから

・母平均の差の信頼空間

$N(\mu_1,\sigma_1^2), N(\mu_2,\sigma_2^2)$の2つの母分散が等しい場合

標本 $(X_1,X_2\cdots X_m)$　標本 $(Y_1,Y_2\cdots Y_n)$

信頼係数$1-\alpha$の信頼区間は

$$\left[\bar{X}-\bar{Y}-t_{\frac{\alpha}{2}}(m+n-2)\,s\sqrt{\frac{1}{m}+\frac{1}{n}},\ \bar{X}-\bar{Y}+t_{\frac{\alpha}{2}}(m+n-2)\,s\sqrt{\frac{1}{m}+\frac{1}{n}}\right]$$

(2つの母分散が等しいと仮定できない場合 ウェルチの近似法を用いて信頼区間を得る)

Date . . No.

・二項分布の信頼区間

信頼係数$1-\alpha$の信頼区間は

$$\left[\hat{p}-z_{\frac{\alpha}{2}}\sqrt{\hat{p}(1-\hat{p})/n},\ \hat{p}+z_{\frac{\alpha}{2}}\sqrt{\hat{p}(1-\hat{p})/n}\right]$$

・ポアソン分布の信頼区間

信頼係数$1-\alpha$の信頼区間は

$$\left[\hat{\lambda}-z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\lambda}}{n}},\ \hat{\lambda}+z_{\frac{\alpha}{2}}\sqrt{\frac{\hat{\lambda}}{n}}\right]$$