『都市工学の数理　基礎編』（浅見泰司）補足メモ

ページ番号は p.で、行番号は l.で表す。なお、－付きの行番号は下から数えることを意味するものとする。

p.3、l.-3～-2 : 正確には加算無限個であり、整数と 1 対 1 に対応づけられることを意味する。
p.6、l.13 :「統計値」→「特性値」。
p.6、l.16 : 平均値 : 期待値とも呼ぶ。
p.8～p.10 : 1-4 節と 1-5 節は順序を逆にする方が説明上わかりやすいかもしれない。
p.10、l.-6～3 行 : 正確には、「ある事象が起きる確率 p が一定の試行（**ベルヌーイ試行**という）を n 回独立に繰り返す場合に，その事象が起きる数の確率を与えるのが**二項分布**（binomial distribution）である．」となる。
p.10、l.-1 : より厳密には、確率変数を大文字、実現値を小文字で表すとすると、

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x} \qquad (x=0,1,...,n)$$

となる。
p.11、l.9 : 後で出てくるベータ関数との区別を明確にするために、

$$x \sim Bin(b,p)$$

と表記する方が良いかもしれない。なお、～はその前の確率変数がその後の分布に従うことを意味する記号である。
p.16、l.13 :「は，n を大きくすれば，平均 0，分散 1 の正規分布に近づく．」→「の分布関数は，n を大きくすれば，平均 0，分散 1 の正規分布の分布関数に近づく．」。
p.16、l.19 :「コーシー分布は平均値や分散がない！」平均値を計算しようとすると、正の側も負の側も発散してしまい定まらない。そのため平均値が存在しない。分散は平均値が無ければ計算できず、やはり存在しない。よって、コーシー分布から得られた標本があっても、その平均値でコーシー分布の中央値を推定することができない。（p.19 の(1)コーシー分布の説明としても同様。）
p.17、l.9 :「変数」→「確率変数」。確率変数を大文字、実現値を小文字で表している。
p.18、l.8 : 上記の二項分布に関する表現方法を用いれば、「$X \sim B(n_1,p)$，$Y \sim B(n_2,p)$ ならば $X+Y \sim B(n_1+n_2,p)$」→「$X \sim Bin(n_1,p)$，$Y \sim Bin(n_2,p)$ ならば $X+Y \sim Bin(n_1+n_2,p)$」となる。
p.19、l.4 : 正確には、「$Z=X+Y$」→「X と Y が独立な場合に，$Z=X+Y$」となる。
p.21、l.9 :「カイ二乗検定で使う分布である．」→「カイ二乗検定で使う分布である．その際には，n は自由度である．」。
p.21、l.-2 :「その際には，n は自由度である．」→「その際には，n は自由度である． n=1 の場合は，コーシー分布と一致する．$n \to \infty$の場合は，正規分布となる．」。
p.24、l.-7 :「**1-11　分布のあてはめ**」→「**1-11　分布の推定**」。
p.25、(3)kernel 法 : カーネル密度推定（kernel density estimation）法とも呼ばれる。$\{x_i: i=1,...,n\}$

をある確率密度関数fに従う独立なn個の標本のときに、その確率密度関数fを

$$\hat{f}_h(x)=\frac{1}{nh}\sum_{i=1}^{n}k\left(\frac{x-x_i}{h}\right)$$

という関数で推定する方法である。$k()$はカーネル関数、hはバンド幅と呼ばれる。このカーネル関数として、三角分布、正規分布、負の二乗分布などが使われる。たとえば、正規分布の場合は、標準正規分布を使う。その場合は、

$$k(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

となる。

p.29、l.-11 :「もう一つよく使われる記号にVがある．これは分散を求める記号である．」→「もう一つよく使われる記号として，分散を求める記号Vがある．」。

p.31、l.-1 :「X_1〜X_n」→「$X_1,\cdots,X_n$」。

p.33、l.-7 : 正確には、「クラスAの数学の点数とクラスBの数学の点数が、期待値が同じ分布から取られた標本と見なせるか」の検定である。見なせない場合には、有意に異なるということになる。

p.39、l.1 : Tは自由度n-1のt分布に従う確率変数である。

p.47、l.7 :「20.4」→「21.6」（2カ所)、「8.5＜」→「9.0＞」

p.59、l.8 :「これまでの検定では」→「2-6節までの検定では」。

p.107、l.-8 :「**動的計画**問題（dynamic programming）」→「**最適制御**（optimal control）」。

p.108、l.14-15 :「最大値ではなく極大値の条件」→「最大値ではなく極大値の必要条件」。正確には、停留値の条件になっているだけで、後述しているように、極大値か、極小値か、鞍点（極大値でも極小値でもなく、たまたま傾きが0となるような点）である可能性がある。

p.124、l.-3 :「$\partial L(\mathbf{x}^*,\lambda^*,0)/\partial\varepsilon=\lambda^*$」→「$\dfrac{\partial L(\mathbf{x}^*,\lambda^*,0)}{\partial\varepsilon}=\lambda^*$」。

p.131、l.-6 : **Karush-Kuhn-Tucker (KKT) 条件** : 例えば、$y\geq 0$というyに関する符号条件が制約条件になり場合は、２行目の一つの等式と２つの不等式条件の代わりに、$\dfrac{\partial L}{\partial y}=0$という一つの等式条件だけに置き換えれば良い。

p.132、l.13 : (2)一般の場合 : さらにより一般の場合としては、以下のような問題を考えることができる。

$$\max_{\mathbf{x}} f(\mathbf{x})$$

$$\text{s.t.}\qquad g_j(\mathbf{x})\leq 0\qquad j=1,\cdots,m$$

$$h_k(\mathbf{x})=0\qquad k=1,\cdots,l$$

このKKT条件は以下のようになる。

$$\nabla f(\mathbf{x}) = \sum_{j=1}^{m} \mu_j \nabla g_j(\mathbf{x}) + \sum_{k=1}^{l} \lambda_k \nabla h_k(\mathbf{x})$$

$$g_j(\mathbf{x}) \le 0 \qquad j = 1,\cdots,m$$

$$h_k(\mathbf{x}) = 0 \qquad k = 1,\cdots,l$$

$$\mu_j \ge 0 \qquad j = 1,\cdots,m$$

$$\mu_j g_j(\mathbf{x}) = 0 \qquad j = 1,\cdots,m$$

μ_jとλ_kは KKT 乗数と呼ばれる。本文の条件とこの一般の場合が一致することは、不等号条件の制約条件の関数に-x、-yを入れて、最初の等式条件をそのスラック変数について解いて、他の条件に代入すると確認できる。

p.150、l.4～8 :「を満たすような・・・る.」→「を満たすような式を求めたい．しかし，現実には，正確に一致することはまれなので，その誤差がなるべく「小さく」なるようにする.」。

p.152、l.2 :「また，ε_iは相互に独立な確率変数であるとする.」→「また，ε_iは相互に無相関な確率変数であるとする.」

p.153、l.1 :「推定値」→「推定量」。(推定量は推定された確率変数、推定値は推定された実際の値を意味する。)

p.153、l.6 :「仮定をしていなかった」→「分布形を特定していなかった」。

p.153、l.-3～-1 :「標本の誤差をe_iとすると，$e_i = y_i - \hat{a} - \hat{b}x_i$」→「標本の誤差を$\hat{e}_i$とすると，$\hat{e}_i = y_i - \hat{a} - \hat{b}x_i$」。

p.154、l.1 :「$\hat{a}$や$\hat{b}$は」→「$\hat{e}$，$\hat{a}$，$\hat{b}$は」。

p.154、l.2～3 :「この分の２つの自由度がこの計算で失われているために,」この説明はやや感覚的なものであり、より正確には、不偏分散を計算すると、その下の式となる。

p.154、l.6～8 : 真の分散値を標本で推定される不偏推定量で置き換えたときに、t 分布に従う。すなわち、$E(s^2) = \sigma^2$が成り立つようにする。

p.157、l.6 : 尤度とは、誤差分布を仮定したときに、標本の出現する確率密度の積であり、それは、誤差分布の分布パラメータの関数（尤度関数という）として定義される。尤度関数が最大になる分布パラメータの推定値を最尤推定値という。真の分布を推定するときにしばしば用いられる方法である。

p.157、l.-3～-1：「標本の誤差 e_i は $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}}$ 」→「標本の誤差　$\hat{e}_i$ は $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}}$ 」。

p.158、l.1：「$e_i =$ 」→「$\hat{e}_i =$ 」。

p.158、l.2：「$\hat{a}$ や $\hat{b}_j$ は」→「$\hat{e}_i$、$\hat{a}$、$\hat{b}_j$ は」

p.158、l.7：ここで、$E(s^2) = \sigma^2$ が成り立つ。

p.158、l.9～10：「すなわち,」及びその次の行の式を削除。この式は 1 変数の時のみに成立するものであり、多変数の場合には、このように表すことができない。

p.162、注 5-4：$\mathbf{P_X} = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$ と定義すると、$\mathbf{P_X}^{\mathrm{T}} = \mathbf{P_X}$、$\mathbf{P_X}^{\mathrm{T}}\mathbf{P_X} = \mathbf{P_X}$ である、ここの導出の式は下記のようにコンパクトに書くことができる。なお、下記で、残差も推定値なので^を付けて表記している。

$$\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\mathbf{b}} = \mathbf{y} - \mathbf{P_X}\mathbf{y} = (\mathbf{I} - \mathbf{P_X})\mathbf{y}$$

$$\begin{aligned}\hat{\mathbf{e}}^{\mathrm{T}}\hat{\mathbf{e}} &= \mathbf{y}^{\mathrm{T}}(\mathbf{I} - \mathbf{P_X})^{\mathrm{T}}(\mathbf{I} - \mathbf{P_X})\mathbf{y} = \mathbf{y}^{\mathrm{T}}(\mathbf{I} - \mathbf{P_X}^{\mathrm{T}} - \mathbf{P_X} + \mathbf{P_X}^{\mathrm{T}}\mathbf{P_X})\mathbf{y} \\ &= \mathbf{y}^{\mathrm{T}}(\mathbf{I} - \mathbf{P_X} - \mathbf{P_X} + \mathbf{P_X})\mathbf{y} = \mathbf{y}^{\mathrm{T}}(\mathbf{I} - \mathbf{P_X})\mathbf{y}\end{aligned}$$

$$\begin{aligned}E(\hat{\mathbf{e}}^{\mathrm{T}}\hat{\mathbf{e}}) &= E[\mathbf{y}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{P_X})\mathbf{y}] = E[\mathrm{tr}(\mathbf{y}^{\mathrm{T}}(\mathbf{I}_n - \mathbf{P_X})\mathbf{y})] = E[\mathrm{tr}([(\mathbf{I}_n - \mathbf{P_X})\mathbf{y}\mathbf{y}^{\mathrm{T}})] \\ &= \mathrm{tr}((\mathbf{I}_n - \mathbf{P_X})E[\mathbf{y}\mathbf{y}^{\mathrm{T}}]) = \mathrm{tr}((\mathbf{I}_n - \mathbf{P_X})(\sigma^2\mathbf{I}_n)) = \sigma^2\mathrm{tr}(\mathbf{I}_n - \mathbf{P_X}) \\ &= \sigma^2[\mathrm{tr}(\mathbf{I}_n) - \mathrm{tr}(\mathbf{P_X})] = \sigma^2[n - \mathrm{tr}(\mathbf{P_X})] = \sigma^2[n - \mathrm{tr}(\mathbf{I}_{p+1})] = \sigma^2(n - p - 1)\end{aligned}$$

p.165、l.-1：「被説明変数」→「目的変数」。

p.166、l.-7～-6：「$D_{\mathrm{M}} + D_{\mathrm{F}} = 1$ となるため，独立ではない.」→「$D_{\mathrm{M}} + D_{\mathrm{F}} = 1$ となるため，この２つのダミー変数の回帰係数は不定となってしまう.」。

p.168、l.-9：「**5-5　重回帰分析の幾何学的解釈**」→「**5-5　重回帰分析における相関係数の幾何学的解釈**」

p.170、l.8～9：「重決定係数」→「重決定係数（決定係数と同じだが，MS-Excel では重決定 R^2 と表記される）」

謝辞

補足メモを執筆するにあたり、東京大学大学院新領域創成科学研究科の本田利器先生、東京大学空間情報科学研究センターの丸山祐造先生、東京大学大学院工学系研究科の片山浩之先生に貴重なコメントをいただいた。記して、謝意を表する。