工学ためのデータサイエンス入門
ーフリーな統計環境 R を用いたデータ解析ー

正誤表 (2005.11.26 現在)

十分注意したつもりでしたが、それでも既に多くの誤植、ミスが見付かっております。以下に現在判明している箇所をメモしておきますので、ご参考にして頂ければ幸いです。読者の皆様に著者一同お詫びを申し上げます. これまでわざわざ御連絡頂きました方々(青木繁伸、越智義道、桜井裕仁、佐藤学、竹澤邦夫、永田靖、平野勝臣、藤井徹、安松勲、元山斉、新井宏嘉、安松勲、窪田央一様) にお礼申し上げます。

6 頁	R では標本分散の定義の既定は不偏標本分散 V_x^2 (56 頁の定義式参照) です。したがって 6 頁の「R なら (2)」中の `var(x)` 及び `sd(x)` はそれぞれ (S_x^2 及び S_x ではなく) V_x^2 と V_x を計算します
6 頁下から 2 行	『5 数要約』を『5 数要約 (発生乱数毎に異なるので注意)』に変えて下さい
8 頁	「R なら (4)」中の下から 2 行目の先頭にプロンプト『>』を加える
11 頁 18 行	(およびその脚注、更に 247 頁文献 [3] 中)『青木繁氏』は『青木繁伸氏』です。お詫びします
12 頁下 2 行	(「R なら (5)」) 図 1.5 と同じ図を得るためには `density(x)` を、`density(x, bw="SJ")` とする必要があります
14 頁	脚注引用文献『渡辺 [23]』は『渡部他 [23]』です

23 頁	コラム「円周率はなぜ π か」中で、円周率を表す記号として π を使いだしたのはオイラーであると書きましたが、これは不正確でした。確かに π が円周率を表す記号として定着したのは、オイラーの 1748 年出版の広く読まれた著書「無限小解析入門」で使われたのがきっかけだったらしいですが、既に 1647 年に π を円周率を表す記号として使った例があるそうです。オイラーも当初は円周率を c とか p とか書いていたそうです。詳しくは「π – 魅力の数」、ジャン＝ポール・ドゥラエ著、畑政義訳、朝倉書店 (2003) を見てください。
23 頁	コラム「円周率はなぜ π か」中で π は『$\pi\epsilon\rho\iota'\mu\epsilon\tau\rho os$ (英語の perimeter) の頭文字 π がその語源』と書きましたが、これは『$\pi\epsilon\rho\iota'\mu\epsilon\tau\rho o\varsigma$ (perimetros、英語の perimeter) の頭文字 π がその語源』に訂正お願いします。ここで最後のギリシャ文字「ς」は σ の変形体ですが、24 頁の一覧表から記載されておりませんので、追加願います。更に、一覧表には ϕ の変形体「φ」が抜けております
25 頁	「1 章の問題の第 4 問」中の『相関係数が 1 になる』は『相関係数の絶対値が 1 になる』です
25 頁問 5	『内閣府総務省』は『総務省統計局』です。また URL は `http://www.stat.go.jp/index.htm` です。
29 頁	14行目「る. 複数の事象 $A_1, A_2, \cdots, A_n$ は関係」を「る. 複数の事象 $A_1, A_2, \cdots, A_n$ は，任意の一部分 $A_{i_1}, A_{i_2}, \cdots, A_{i_m}$, $1 \leq i_1 < i_2 < \cdots < i_m \leq n$ に対して」に変更して下さい.
30 頁	式を「$\boldsymbol{P}\{A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_m}\} = \boldsymbol{P}\{A_{i_1}\}\boldsymbol{P}\{A_{i_2}\} \cdots \boldsymbol{P}\{A_{i_m}\}$ (事象の独立性)」に変更して下さい.
31 頁 11 行	(及び 33 頁 8 行) 式中の『$\sum_i p(x) = 1$』は『$\sum_i p(i) = 1$』です
31 頁 15 行	『quantail』は『quantile』です
¿ 33 頁 13 行	『$\sum_i (a_i - \mu)^2 p(a_i)$』は『$\sum_i (a_i - \mu)^2 p(i)$』です
34 頁下 4 行	『cavarinace matrix』は『covariance matrix』です
37 頁下 9 行	『χ_n』を『χ_n^2』に変えて下さい
38 頁 8 行	『確率変数 X, Y の比 X/Y の確率分布』は『確率変数 X, Y に対する比 $(X/m)/(Y/n)$ の確率分布』です
39 頁 12 行	多変量正規分布の密度関数の式の指数関数部分に 1/2 をいれる必要があります： $$f(x) = \frac{1}{2\pi^{p/2}\|\Sigma\|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$
39 頁	10,13,19 頁にある 4 つのベクトル X, x, μ は縦ベクトルと考えていますので、転置記号 ^{T} を付ける必要があります。例えば『$X = (X_1, X_2, \ldots, X_p)$』は『$X = (X_1, X_2, \ldots, X_p)^T$』とします
39 頁 14,20 行	『$x = (x_1, x_2, \ldots, x_p)$』, 『$\mu = (\mu_1, \mu_2, \ldots, \mu_p)$』, 『$\mu = (\mu, \mu, \ldots, \mu)$』をそれぞれ 『$x = (x_1, x_2, \ldots, x_p)^T$』, 『$\mu = (\mu_1, \mu_2, \ldots, \mu_p)^T$』, 『$\mu = (\mu, \mu, \ldots, \mu)^T$』に変えて下さい。
39 頁下 8 行	『Σ_{ij}』は『σ_{ij}』です
45 頁 9 行	『`.Machine$double.neg.eps`』は『`.Machine$double.eps`』です

47 頁	定理 (2 項分布のポアソン近似) 中の『$k = 0, 1, 2, \cdots$』は『$i = 0, 1, 2, \cdots$』です
55 頁下 2,3 行	(および 56 頁下 2 行) 『S^2』は『S_x^2』です
56 頁下 5 行	一番右の式中の項『$(x_i - \bar{x})$』は『$(x_i - \bar{x})^2$』です
58 頁	表 2.13 中の最後の三つの列はそれぞれ『身長 16、人数 394』を『身長 75、人数 16』、『身長 5、人数 669』を『身長 76、人数 5』、『身長 2、人数 990』を『身長 77、人数 2』、に変更願います
58 頁数式 2 行	$\frac{X-4}{\sqrt{3/8}}$ は $\frac{X-4}{\sqrt{8/3}}$ です。
61 頁 1 行	『$p(i) \simeq \frac{1}{i^a}$』は『$p(i) \propto \frac{1}{i^a}$』です
61 頁 5 行	『100』は『1000』です
64 頁 19 行	Hastings の近似式の変数範囲『$X \geq 0$』は『$x \geq 0$』です
66 頁 ~~6,7~~ 5,6 行	R 1.9 より階乗関数を計算する専用関数が導入されましたので，以下のように書き換えます。「R では階乗関数 `x!` は `factorial(x)` で計算される．これは単に `gamma(x+1)` を計算しているだけで，`x` は 0 および負の整数以外の実数でも良い.」
66 頁 9 行	「`lgamma, lchoose`」を「`lgamma, lfactorial, lchoose`」に変更する.
注意	61 頁 1 行の Zipf 分布の定義では指数 a は $a > 1$ でないと総和不能 (有限な和を持たない) です。67 頁 10 行の一般化された Zipf 分布の定義式でも $a \leq 1$ の場合は必ずしも総和可能ではありませんが、ここで考えている例では $c < 1$ なので、$a < 1$ にもかかわらず総和可能です
75 頁下 2 行	$\sum_{i=0}^{n}$ は $\sum_{i=1}^{n}$ です。
76 頁 8,9,12,13 行	$\sum_{i=0}^{n}$ は $\sum_{i=1}^{n}$ です。
76 頁 8 行	『$+\frac{(x_i-\mu)^2}{2s}$』は『$-\frac{(x_i-\mu)^2}{2s}$』です。また下 4,5 行の式中の分母の『$2s$』は『s』です
80 頁 9 行	『帰無仮説らしくなくても』は『帰無仮説らしくても』です
80 頁 10,11 行	『5%』は『1%』、『1%』は『5%』です (それぞれ 2 箇所)
86 頁 6 行	Welch 検定統計量の定義式の分母『$V_x^2/m + V_y^2/n$』は『$\sqrt{V_x^2/m + V_y^2/n}$』に訂正願います。また 8 行の Wlech 検定の自由度の定義式は**分母・分子がひっくり返っています**ので、訂正願います
87 頁 10 行	『$\mu_x \neq \mu_y$』は『$\mu_x = \mu_y$』です
88 頁 3 行	『両標本の母集団分散が異なる $\sigma_1^2 \neq \sigma_2^2$ という帰無仮説』は『両標本の母集団分散が等しい $\sigma_1^2 = \sigma_2^2$ という帰無仮説』です
88 頁	「R ならば (13)」中の『`extra[group==1]`』, 『`extra[group==2]`』をそれぞれ『`x <- extra[group==1]`』, 『`y <- extra[group==2]`』に訂正願います。また『`x <- sleep`』は不要です
89 頁下 3 行	の『`var.equal=FALSE`』は『`var.equal=TRUE`』、下 2 行の『`var.equal=TRUE`』は『`var.equal=FALSE`』です
94 頁 5 行	『サイズ 4807 個』は『サイズ 480 個』です
96 頁表	表中のハーストのデータに対する p 値は 0.7408 です。
105 頁下 4 行	『$f_\theta(x)$ が x の』は『$f_\theta(x)$ が θ の』です

107 頁 11 行	傾きの推定値式分子中の項『$(y_i - \bar{x})$』は『$(y_i - \bar{y})$』です
113 頁 16 行	予測誤差分散式中の第二項分母の『S_x^2』は『nS_x^2』です
114 頁 1 行	予測値の誤差分散推定式中第二項分母の『S_x^2』は『nS_x^2』です
114 頁下 2 行	式『$\sqrt{\frac{1}{n} + \frac{(x'-\bar{x})^2}{nS_x^2}}$』を『$\sqrt{1 + \frac{1}{n} + \frac{(x'-\bar{x})^2}{nS_x^2}}$』に変えてください
116 頁 3 行	『`optden = carb`』は『`optden ~ carb`』です
116 頁下 8 行	『`summary(fm1)`』は『`summary(fm)`』です
118 頁最初の表	表の見出しの『傾き α』を『切片 β』に、『切片 β』を『傾き α』に交換してください。
119 頁下 1 行	120 頁のグラフと同じ物を得るには『matplot(new$carb, cbind(dp,dc), lty=c(1,2,2,3,3), type="l")』を『matplot(new$carb, cbind(dp,dc), lty=1, col=c("black","black","black","black","blue","blue"),type='l')』に交換してください
119 頁下 8 行	「R なら (17)」中の 3,7 行の『`predict(d, new,...)`』は『`predict(fm, new,...)`』です
122 頁下 4 行	の『$(2\pi)^{-1}$』は『$(2\pi\sigma^2)^{-1}$』です
128 頁 4 行	式中の項『$(y - X\theta)(y - X\theta)^T$』は『$(y - X\theta)^T(y - X\theta)$』です
134 頁下 7 行	『`Ed:Me`』は『`Ed:In`』です
145 頁下 8 行	『符号が大きければ』は『パラメータ推定値が大きければ』です
148 頁下 7 行	『`agregate`』は『`aggregate`』です
157 頁 4 行	『$\hat{\varepsilon}_{ij} = y_{ij} - \hat{\alpha}_i - \hat{\beta}_j$』は『$\hat{\varepsilon}_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$』です
157 頁 11 行	式中の『χ_A^2』は『χ_M^2』です
175 頁	第 7 章前書き中の『目的変数が説明変数の線形関数でない』は不正確で『母回帰式が未知パラメータの線形関数でない』に訂正願います。
189 頁最下行	『体重 W(ワット) と基礎代謝量 E(kg)』は『体重 W(kg) と基礎代謝量 E(ワット)』です
189 頁	コラム中の Boyd 公式を次のように訂正願います (ただしここでの体重 W はグラム単位) $$0.0003207 \times H^{0.3} \times W^{0.7285 - 0.0188\log_{10}(W)}$$ ちなみに私 (間瀬、体重 59kg、身長 167cm) では Boyd 式 $A = 1.661775m^2$、Gehan-George 式 $A = 1.653885m^2$、Mosteller 式 $A = 1.654371m^2$、Haycock 式 $A = 1.653596m^2$ で、総体表面積は約 $1.3m$ 四方となります (座蒲団カバーに使える位?)
190 頁 24 行	の『大局的最小値は』は『大局的最小値を与えるパラメータ値は』です
201 頁下 3 行	『`n=2,...,50`』は『`n=2,...,20`』です
202 頁下 2 行	(および 203 頁上 1 行) 式中の『$(n-k+1)^{2(n-k)}$』は『$(n-k+1)n^{2(n-k)}$』です

209 頁	囲み記事中の R コードの下 1,2,3 行の先頭にある R のプロンプト『>』は不要です
215 頁下 5 行	『`i0 <- sample(1, 1:n)`』は『`i0 <- sample(1:n, 1)`』です。なおここで紹介したコードはアルゴリズムを忠実にプログラム化したものですが、実行速度は遅く、大量のデータの処理はできません。この正誤表の末尾に S 言語の開発者であり、現在は R の開発者チームの一員でもある J. Chambers の本に紹介されている洗練されたコードを参考のために紹介します。
220 頁 10 行	『`x2 <- round(x0)`』は『`x2 <- round(x1)`』です
225 頁 4 行	『一様乱数 x, y を発生し』は『一様乱数 x, y を用いて点 $(2x-1, 2y-1)$ を発生し』です
225 頁 5 行	『$x = \cos(2\pi\theta), y = \sin(2\pi\theta)$』は『$x = r\cos(2\pi\theta), y = r\sin(2\pi\theta)$』です
231 頁 15 行	『R 例示用の』は『R は例示用の』です
231 頁	脚注中の『Tck/Tk』は『Tcl/Tk』です
233 頁 11 行	『岡田昌史』は『岡田昌史氏』です。お詫びします
223 頁下 8 行	『層化無作為二段階抽出法』は『層化二段階無作為抽出法』が普通の呼び名のようです。
234 頁 12,13 行	『そして RjpWiki のメンバーが計画中の本 (刊行時期未定) が参考になるであろう。』を『岡田昌文編「The R Book」[6]、および舟尾暢男著「The R tips—データ解析環境 R の基本技・グラフィックス活用集」[7] が参考になるであろう。』に変更して下さい
236 頁下 2 行	『CVS』は『CSV』(Comma Separated Values) です
241 頁 7 行	『$\mu\sum_{i=1}^{\infty} e^{-\mu}\frac{\mu^{i-1}}{(i-1)!}$』は『$\mu\sum_{i=1}^{\infty} i\times e^{-\mu}\frac{\mu^{i-1}}{(i-1)!}$』です
243 頁下 9 行	『$-2\sum_{i=1}^{n}\hat{\epsilon}(y_i-\bar{y})$』は『$+2\sum_{i=1}^{n}\hat{\epsilon}(\hat{y}_i-\bar{y})$』です
243 頁下 7 行	『$y_i-\bar{y}$』は『$\hat{y}_i-\bar{y}$』です
245 頁 13 行	『$\frac{1}{s^2}$』は『$\frac{1}{2s^2}$』です
248 頁	文献 [6] 中の『船尾暢男氏』は『舟尾暢男氏』に訂正願います。お詫びします
248 頁	文献 [7]『数理統計学 (稲垣宣生著)』は改訂版が 2003 年に出ています
248 頁	参考文献 [6] を、以下の様に変更して下さい。『舟尾暢男著「The R tips – データ解析環境 R の基本技・グラフィックス活用集」、九天社 (2005)。なお、この本の前身は URL `http://cse.naro.affrc.go.jp/takezawa/index2.html` で公開されている。』
249 頁	文献 [23] の著者名『渡辺』は『渡部』です。またこの本の題名を『探索的データ解析入門 –データの構造を探る–』に訂正願います
249 頁	文献一件追加.『岡田昌史編「The R Book」, 九天社 (2004). R のインストール、R を使った統計手法、パッケージの使い方等を紹介.』

66 頁コラム中の例 $\binom{10000}{100}$ は例のように計算しなくても、直接計算可能でした。したがって例としてはたとえば $\binom{10000}{1000}$ を考えるのが適切でした。以下のように訂正願います

```
> choose(10000,100)                    # これは直接計算可能
[1] 6.520847e+241
> choose(10000,1000)                   # これは大きすぎて無限大とされる
[1] Inf
> x <- lchoose(10000,1000)/log(10) # 自然対数値を計算、更に常用対数
に直す
> x
[1] 1409.941                           # 指数部は 1409
> 10^(x-1409)                          # 仮数部を計算する
[1] 8.733076
```

つまり $\binom{10000}{1000}$ は 8.733076×10^{1409} となる。

198 頁のコード中の上から 6 行目の `x <- c(x,wa)` は `x[i] <- wa` とするほうが効率的です。また 199 頁, 214 頁のコードも効率性のために次のように書き換えたいと思います：

```
birthday <- function(n){  # 199 頁
   rep <- 1000
   x <- numeric(rep)
   for (i in 1:rep){
      u <- sample(365, n, replace=TRUE)
      u2 <- outer(u, u, "==")
      x[i] <- length(which(u2==TRUE))}
   r <- length(x[x>n])
   return(r/rep)
}

jiko <- function(lm1, lm2){ # 214 頁
   rep <- 10000
   rel <- numeric(rep)
   for (i in 1:rep){
      n <- rpois(1, lm1)
      rel[i] <- sum(rpois(n, lm2))}
   return(rel)
}
```

215 頁のクイックソートアルゴリズムの J. Chambers によるプログラム例 (少し修正) を参考のため紹介します (「データによるプログラミング」垂水他訳、森北出版、2002、68 頁)。これは R 初心者には理解しにくいプログラムですが、R 言語の特徴をうまく利用した優れたプログラムです。なお、R の組込みソート関数 `sort()` は内部的に C 言語で記述したクイックソートプログラムを使った高速な関数です。

```
quicksort <- function (x) {
   if (length(x) <= 1) return(x)
   if (length(x) == 2 && x[1] <= x[2] ) return(x)
   if (length(x) == 2 && x[1] >  x[2] ) return(x[2:1])
   fence = sample(x,1)
```

```
    return( c(quicksort(x[x < fence]), x[x == fence],
            quicksort(x[x > fence])) )
}
```

時間をはかってみると次のようになりました（結果は当然すべて同一です)。

```
> system.time(quick(10000:1))        # 修正済みの quick 関数
[1] 4.49 0.04 4.53 0.00 0.00         # 実行時間 4.49 秒
> system.time(quicksort(10000:1) )   # Chambers の quicksort 関数
[1] 0.5 0.0 0.5 0.0 0.0              # 実行時間 0.5 秒
> system.time(sort(10000:1))         # R の組込みソート関数 sort()
[1] 0.01 0.00 0.00 0.00 0.00         # 実行時間 0.01 秒
```

139 頁の内容は担当著者の理解不足から誤った説明になっており、全面的な書き換えが必要です。また leverage(梃子) という用語は既に**梃子比**という訳語が定着していますので、併せて訂正させて頂きます。お詫びします。とりあえず、次のように節タイトルを含め (関連脚注は抹消) 変更し、識者の御批判を頂きたいと思います (竹澤邦夫様からはこの件に関し何度も丁寧なコメントを頂き感謝しております)。

5.4.2 影響力の大きなデータのチェック – 梃子比

データの説明変数のパターンが、当てはめ回帰式に対する、ある特定の目的変数の値の影響 (influence) を強調することがある。H を回帰モデル式のハット行列とすると，残差ベクトルは $\widehat{\epsilon} = (I - H)\epsilon$ と表される．各誤差が等分散性 $\mathrm{Var}\{\epsilon\} = \sigma^2 I$ という仮定が正しければ，残差ベクトルの分散共分散行列は $\mathrm{Var}\{(I - H)\widehat{\epsilon}\} = (I - H)\sigma^2$ となる．特に誤差が互いに独立であっても，残差は一般に互いに相関を持つようになる．行列 H の第 (i, i) 成分を h_i と置くと，個々の残差の推定値の分散は $\mathrm{Var}\{\widehat{\epsilon}_i\} = (1 - h_i)\sigma^2$ となる．これより，h_i が 1 に近ければ $\mathrm{Var}\{\widehat{\epsilon}_i\}$ は本来の分散値 σ^2 よりも小さくなり、$\widehat{y}_i \simeq y_i$ となることが予想される。値 $\{h_i\}$ を**梃子比** (てこひ，leverage) と呼ぶ．つまり，予測式は他のデータよりも、i 番目のデータにより「近付く」ようになる。このことは $H = (h_{ij})$ とすれば関係 (127 頁の予測値ベクトル公式)

$$\hat{y}_i = h_i y_i + \sum_{j \neq i} h_{ij} y_j$$

が成り立つことから、梃子比 h_i が大きなデータ番号 i に対しては、説明変数 y_i の値がその分増幅されて予測値 $\hat{y}_i$ に反映する (梃子比の名前のいわれ) ことからも了解される。更に h_i は、x_i のその中心 $\bar{x}$ からのある種の距離に関係しており、極端な x_i の値は対応する梃子比を大きくする傾向がある。

一般的性質として，全ての i で $h_i \leq 1$, $\sum_i h_i = p$ (ここで p は説明変数の数で、もし定数項があれば、それを含めた数である)、$h_i \geq 1/n$ が成り立つので，h_i の「平均的な値」は $(p + 1)/n$ と見積もることができ，例えば $2(p + 1)/n$ より大きな h_i は「特異」と判断できる．実践的見地からは、$0.2 < h_i \leq 0.5$ ならば「対応データは危険」、そして $0.5 < h_i$ ならば「対応データは解析からは除外」することが提案されている。但し、梃子比が大きいことだけを以て、対応する y_i が外れ値であると判断することは困難である。間違いなくいえることは、梃子比

が大きめで、したがって予測誤差の分散が小さのはずなのに、実際の残差 $y_i - \hat{y}_i$ が他に比べて大きめならば、そのデータ y_i は「挙動不審」という事である。

例えば、R の基本パッケージの組込みデータである「Anscombe の四つ組データ」は、全く同じ線形回帰式を与える、全く雰囲気の異なる 4 種類の教訓用 (人工) データであるが、次の図の左下 (第 3 データ) から、明らかに第 3 データは外れ値と判断すべきであるが、その梃子比は 0.236 であり、一方同じかそれ以上の比を持つ第 6,8,11 番目のデータ (梃子比はそれぞれ 0.318,0.318,0.236) は問題があるとは見えない。また同じ図の左下の第 4 データでは、梃子比は第 8 データを除き全て 0.1 であるが、第 8 データの梃子比は可能な最大値 1 を持つ。実際第 8 データ (一番左端のデータ) に当てはめ式が振り回される (影響が大) ことは、図から明らかであろう。但し、だからといって必ずしも第 8 データが外れ値であると即断できない。ちなみに、Anscombe の第 1,2,3 データは説明変数が全く同じであり、したがって全く同じ梃子比を持つ事を注意しておこう。

R で梃子値を吟味するコードは，例えば以下のようになる．この例では，梃子比が大きめのデータは 2 つであった。

R なら (24): 回帰診断 (梃子比の吟味)

```
> data(swiss)                              # 再び swiss データを使用
> d1 <- lm(Fertility ~ ., data=swiss) # 相互作用項無しのモデル
> d2 <- step(d1)                           # 変数選択
> X <- model.matrix(d2)                    # モデルの計画行列
> lev <- hat(X)              # 梃子比ベクトル (ハット行列の対角成分)
> cx <- 2*sum(lev)/dim(swiss)[1]           # 警戒水準値 2*5/47 を計算
> cx
[1] 0.2127660
> names(lev) <- rownames(swiss)            # 梃子比に地域名を対応させる
> lev[lev > cx]                            # 要注意データを表示
   La Vallee V. De Geneve                  # その地域名
   0.3181944    0.4554370                  # その梃子比
> plot(lev)                                # 梃子比ベクトルのプロット
> abline(h=cx)                             # 警戒水準線を重ね描き
```

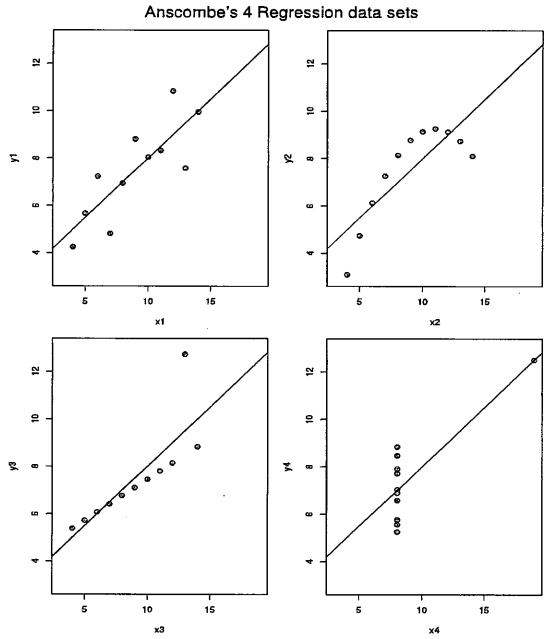

Figure 1: Anscombe の四つ組データの回帰分析結果のプロット

工学ためのデータサイエンス入門
ーフリーな統計環境 R を用いたデータ解析ー

第 4 版用正誤表 (2007.03.07 現在) : 初版、第 2 版、第 3 版用正誤表は、修正しておりませんので、各版の正誤表をお使いの方は、本正誤表を追加してください。

十分注意したつもりでしたが、それでも既に多くの誤植、ミスが見付かっております。以下に現在判明している箇所をメモしておきますので、ご参考にして頂ければ幸いです。読者の皆様に著者一同お詫びを申し上げますとともに、もし更に間違いと思われる点に気づかれた場合は、ご面倒でも **mase@is.titech.ac.jp** 宛にお知らせくだされば幸いです。

12 頁下から 11 行	『間欠泉の噴出時間間隔』は『間欠泉の噴出継続時間 (分単位)』です。
30 頁 18 行	『$1-F(-x)+F(x-0)=P\{\|X\|\le x\}$』は『$1-F(x-0)+F(-x)=P\{\|X\|\le x\}$』です。
76 頁 3 行	『パラメータ n』は『パラメータ m』です
106 頁	線形単回帰モデルの正規方程式において 2 つの式において $\sum_{i=1}^{n}$ の前に係数 2 が抜けています。
114 頁 3 行	『上側 $100(p/2)\%$ 点』は『上側 $100((1-p)/2)\%$ 点』です。
119 頁 3 行と 4 行	『区間 [0.3,0.6]』は『区間 [0.2,0.8]』です。『$\alpha+\beta x_0$』は『$\alpha x_0+\beta$』です。『$\hat{\alpha}+\hat{\beta}x_0$』は『$\hat{\alpha}x_0+\hat{\beta}$』です。
130 頁 9 行	『 $AIC=2p+\frac{n}{2}+\frac{n}{2}\log(2\pi\widehat{\sigma}^2)=2p+\frac{n}{2}\log(\widehat{\sigma}^2)+$ 定数項』は『$AIC=2(p+1)+n\log\widehat{\sigma}^2+n+n\log(2\pi)$』です。
130 頁下から 2 行	『パラメータ数に σ^2 の分が加算されていないが，AIC 統計量は相対的な大小のみが意味を持つため，誤差の構造が同一である限り，無視してもよい.』は『パラメータ数が p ではなく $p+1$ となっているのは，誤差分散パラメータ　σ^2 を加えているからである.』です。
137 頁下から 3 行	『 $\widehat{\epsilon}=(I-H)\epsilon$』は『$\widehat{\epsilon}=(I-H)y$』です。
139 頁下から 9 行	『 分散が小さの』は『分散が小さい』です。
145 頁下から 11 行	『各データの分散の自乗を重みとした 』は『各データの分散の自乗の逆数を重みとした』です。

157 頁下から 8 行	『 は自由度 $(s-1)(t-1)$ のカイ自乗分布に従うことが知られており，$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{(s-1)(t-1)}\sum_i\sum_j \hat{\epsilon}_{ij}^2 \\ &= \frac{\sigma^2}{(s-1)(t-1)}\chi_e^2\end{aligned}$$ が σ^2 の不偏推定量となる.』は『は自由度 $(s-1)(t-1)$ のカイ自乗分布に従うことが知られており，カイ自乗分布の期待値はその自由度であるから，χ_e^2 に $\sigma^2/\{(s-1)(t-1)\}$ をかけると σ^2 の不偏推定量 $$\hat{\sigma}^2 = \frac{\sigma^2}{(s-1)(t-1)}\chi_e^2 = \frac{1}{(s-1)(t-1)}\sum_i\sum_j \hat{\epsilon}_{ij}^2$$ を得る.』です。
184 頁 1 行	『$\log_{10}(V) = \alpha - t\beta\exp(-\gamma/T) + \epsilon$』は『$\log_{10}(V) = \alpha - t\beta\exp(-\gamma/(T+273.16)) + \epsilon$』です。
184 頁 10 行	『R なら (42):最適解への収束状況の出力』を以下に変更する。

```
> fm <- nls(log10(y) ~ a - b*x1*exp(-c/(x2+273.16)), data=Nelson,
           start=c(a=1.13,b=6.375e+11,c=17065),trace=T)
0.7269386 : 1.1300e+00   6.3750e+11   1.7065e+04   # 初期値
0.722269  : 1.125919e+00 2.878341e+11 1.664299e+04 # 以下途中解
0.7220003 : 1.125479e+00 2.637889e+11 1.659572e+04
0.7213283 : 1.125151e+00 2.473455e+11 1.655993e+04
0.7206689 : 1.124662e+00 2.242644e+11 1.650580e+04
0.7194283 : 1.124181e+00 2.036685e+11 1.644994e+04
0.719196  : 1.124190e+00 2.037838e+11 1.644781e+04
0.719196  : 1.124190e+00 2.037473e+11 1.644772e+04  # 収束値
```

184 頁下から 3 行	『初期値 (1.13,55589,4091)』は『初期値 (1.13,6.375e+11,17065)』です。
186 頁 14 行	『R なら (44):回帰分析要約 **summary(fm)** の出力』を以下に変更する。

```
Formula: log10(y) ~ a - b * x1 * exp(-c/(x2 + 273.16))

Parameters:
   Estimate Std. Error t value Pr(>|t|)
a 1.124e+00  8.244e-03 136.360   <2e-16 ***
b 2.037e+11  4.227e+11   0.482    0.631
c 1.645e+04  1.137e+03  14.469   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.07585 on 125 degrees of freedom
```

186 頁下から 3 行	『非線形回帰曲線 $\log_{10}(y) = \alpha - \beta x_1 \exp(-\gamma/x_2)$』は『非線形回帰曲線 $\log_{10}(y) = \alpha - \beta x_1 \exp(-\gamma/(x_2 + 273.16))$』です。
187 頁 2 行	推定値に関する表を以下のように変更する。

	最小自乗推定値	標準偏差推定値	対応 t 値	その p 値
α	1.124	8.244e-3	136.360	2e-16
β	2.037e+11	4.227e+11	0.482	0.631
γ	1.645e+4	1.137e+3	14.469	2e-16

187 頁 7 行	『(4)Correlation of Parameter Estimates は母数の推定値の相関を示している.』は R2.4.1 では出力されないので削除する。
193 頁 4,5 行	『$(a = 1.13, b = 55589, c = 4091)$』は『$(a = 1.13, b = 6.375e + 11, c = 17065)$』です。
198 頁下から 13 行	『x <- numeric(10000)』は『x <- numeric(rep)』です。
198 頁下から 9 行	『x[i] <- wa』は『 x[i] <- wa}』です。
201 頁下から 10 行	『# b[a] は i 番目の女が選んだ男が選んだ女の番号』は『# b[a] は i 番目の男が選んだ女が選んだ男の番号』です
204 頁下から 10 行	『また，各確率分布の密度・確率関数を求めるには』は『また，たとえば正規分布の密度・確率関数を求めるには』です。
221 頁 2 行	『範囲 [0,M-1) 』は『範囲 [0,M-1]』です。
245 頁下から 11 行	第 7 章の問題 2 の回答の部分を以下のように変更する。

2．色々なやり方が考えられるが Nelson の論文にある一例を示す.

(2) 図 7.3 において，4 種類のデータごとに目測で適当に回帰直線 (実際は曲線) を 4 本引く．この 4 本の直線が $t = 0$ で交わる点が V_0 の初期値の推定値 (仮に 13.5 とする) と考える．すると $\widehat{\alpha} = \log_{10} 13.5 = 1.13$ となる.

(2) モデル式 $\log_{10} V_i = \widehat{\alpha} - t\beta \exp(-\gamma/(T_i + 273.16))$ $(i = 1, 2)$ から γ と β の初期値推定式を作ると次のようになる.

$$\widehat{\gamma} = \frac{(T_1 + 273.16)(T_2 + 273.16)}{T_1 - T_2} \log[\log_{10}(V_0/V_1)/\log_{10}(V_0/V_2)],$$
$$\widehat{\beta} = (1/t) \exp(\gamma/(T_1 + 273.16)) \log_{10}(V_0/V_1)$$

例えば $t = 32, T_1 = 250, T_2 = 275$ での V_1, V_2 を目測で決定 (例えば $V_1 = 9.8, V_2 = 3.27$) する．これを上の式に代入して初期値推定値 $\widehat{\beta} = 6.375e + 11, \widehat{\gamma} = 17065$ を得る.

247 頁 20 行	正しい URL は『http://www.okada.jp.org/RWiki/]』です