

「統計と確率の基礎」(服部哲弥著, 学術図書)

第2版第1刷から第2刷への改訂事項一覧

20080220

p. i l. -1 (第7章以降の練習問題は全て略解をつけることにしたための説明の改訂)

「練習問題には、巻末略解を付けない問題も含めた。問題番号に*をつけてある。また、第6章までの練習問題について、」 「練習問題のうち、第6章までの練習問題について、巻末略解を付けない問題も含めた。問題番号に*をつけてある。また、」

p. ii l. -2 「信州大学」 「首都大学東京」

p. iii l. 2 「沼澤洋平君」 「沼澤洋平君(現三井住友銀行)」

p.18 l.-2 「分布だという点だけである。」 「分布だという点である。」

p.20 l.-2 「, X が実数値をとるならば X の分布 Q は連続散分布になる」 「なる」

(連続散分布という単語は無いのと、連続分布になるためには密度の存在という条件を要するという間違い)

p.21 l.10, -10, -6 「連続値」 「連続型」(3カ所)

(離散値、連続値という表現は確率変数の値域に関する数学用語なのに対して、ここに出てくる連続型確率変数は、その分布が密度を持つ、という条件が加わるので紛れないように訂正する。)

p.22 l.-4 - -1 (間違っていないが、先に行かないと分からない記述のための改訂)

括弧内の2文 トル

p.23 l.8-10 (間違っていないが、あとの改訂のページ数調節のための改訂)

「最初に、...を確認する。」の1文 トル

p.25 l.-1 - p.26 l.8 (記述を圧縮して追加説明のスペースを作るための改訂)

(1) 「平均待ち時間 \bar{t} は一列並びと並列並びで等しいことを説明したが、... すなわち待ち時間の期待値(平均値)は等しい。」の2段落9行を切り取って p.24 l.15 項目5 「以外の条件は変わらないとする。」の直後に次の形で移動する:

「待ち時間を表す確率変数 T は、一列並びと並列並びで異なる。

$$\text{一列並びの場合は } T^{(1)} = \frac{1}{M} \sum_{i=1}^N S_i,$$

(2.19)

$$\text{並列並びの場合は } T^{(2)} = \sum_{i=1}^{N/M} S_{j_i}.$$

ここで $j_1, j_2, \dots, j_{N/M}$ は並列並びについて自分と同じ列に並ぶ客。 N 人が M 個の窓口に分かれて並ぶので自分の前には N/M 人いる。念のため添字を2重にしたが、ここでの分析では客の番号(名前)は重要ではない。

まず、(2.19) と期待値の線形性 (2.14) (2.16) および $\tau = E[S_i]$ から

$$E[T^{(1)}] = E[T^{(2)}] = \frac{N}{M} \tau = \bar{t}$$

すなわち待ち時間の期待値(平均値)は並び方によらない。」

(2) その直後 (p.24 l.16-22 項目5の直後の段落) 「1人の客の処理をするのに平均 τ の処理時間 ... すなわち平均値に関しては Y 駅の主張は正しい。」 「平均値に関しては Y 駅の主張は正しい。」

p.24 l.-3 「前売りを買う人は、」 「仕事に忙しい、まじめな社員が前売りを買うときは、」

p.24 l.-1 「会社に勤務する人などにとっては、」 トル

p.25 l.-2 「もっと精密な考察は勉強が進んだあとの宿題としよう。」 トル

p.26 l.-2 「確率が低いと思われる。」 と 「並んでいられる時間が限られている」の間に以下の段落を挿入：

「(改段落) 並列並びのほうが、早い窓口をめざとく見つけて並ぶことで、自分より先に来たのに遅い窓口で並んでしまった客を追い越せると思うかもしれないが、自分の直前の客がもたついで、後から来て他の窓口で並んだ客が自分を追い越して先に終わる可能性もある。つまり、並列並びのほうが一列並びより極端に待ったり極端に早かったりする両極端が多い。これを『並列並びのほうが待ち時間のばらつきが大きい』と言う。いまは、平均よりも余裕を見て買いに来ているので、極端な場合全員平均どおり処理が終われば必ず自分も間に合う。だからばらつきの目安となる分散が小さい一列並びのほうが目的にかなう。

もし、朝寝坊して平均通りだと間に合わない時刻に列に到着して、一か八か幸運に賭けたい、という賭け事の好きなタイプならば、自分の目の前にいる客の少ない並列並びを好むかもしれない(そして、自分の前の客が偶然全員手際が良いことに賭ける)。でも、ここでは用心深く余裕を持って到着した人の立場で考えている。そういう慎重な人にとっては、ラッキーな人の多い並列並びはそのぶん「割を食う」ので、一列並びのほうが良い。」

p.27 l.7 「インターネットで前売りを予約」 「インターネットで指定席を予約」

p.27 l.13 「角度から議論する。」 「角度から議論することになる。」

p.31 l.-2 「視聴率は決まる。」 「 n 人の中での視聴者の割合は決まる。」

p.36 l.9 系4の証明の直後に次を挿入：

「標本平均は第2章の母平均に対応する用語である。本当は母平均などの母集団の値(母集団特性値)がほしいがそれを標本の値で代用する、というのが統計的推測の出発点である(改段落)」

p.36 l.9 「データの大きさを大きくすると。」 「系4から、データの大きさを大きくすると。」

p.49 l.2 「前期過程」 「前期課程」

p.49 l.3 「後期程」 「後期課程」

p.59 l.11 「長い桁数は意味がない。」 「長い桁数は意味がない。1/6の小数表示は長い桁数も意味があるが、データと比べる場合はデータの桁数以上はいらぬ。」

p.63 l.8 「実測によって常識を覆す」 「実測によって覆す」

p.72 l.5 「有効活用である、ということである。」 「有効活用である(改段落)もう一点、区間(6.2)はデータの大きさ n の情報を含むことも注目値する。点推定では推定値は1つの数字なので結論を得るのに用いた n が分からない。個々のデータはばらついて n が大きければ標本平均の母平均からのずれは小さいことが期待できる。データの大きさ n を情報として提示することも区間推定の重要な意義である(改段落)」

p.73 l.12 「 $\mu = x \pm a\sqrt{vn}$ (99% CL)」 「 $\mu = x \pm a\sqrt{\frac{v}{n}}$ (99% CL)」

p.74 l.-1 - p.75 l.1 「もっと単純に、 $p \approx \bar{X}_n(\omega)$ で近似して $v = \bar{X}_n(\omega)(1 - \bar{X}_n(\omega))$ とする方法もある。どの方法をとっても実用上数字の上では大差ない。」 「もっと単純に、 $v = p(1 - p)$ の部分だけ $p \approx \bar{X}_n(\omega)$ と近似して $v = \bar{X}_n(\omega)(1 - \bar{X}_n(\omega))$ とすると $1/n$ が小さいときの近似を得る。」

p.75 l.8 「一方、この幅は調査対象数（データの大きさ）だけで決まる。」 「(改段落)ところで、この幅は調査対象数（データの大きさ）だけで決まる。」

p.75 l.10 「調査すべき人数は等しい！」 「調査すべき人数は等しい！第3章の(3.1)のところで無作為抽出を説明した際に p は母集団の大きさと無関係な一人一人の番組を見ようという傾向であると考えられることを示唆したが、データは独立同分布確率変数であると決めたときにその視点が組み込まれている（また、独立な確率変数がいくらかでもたくさん用意できると考えるときに、母集団の大きさは無限としたことになる。）たとえば仮にデータの大きさを全人口に近づけても推定区間の幅は0にはならない。」

p.88 l.8 「自由度 n の T 分布といい、 T_n と書く。」 「自由度 n の t 分布といい、 T_n と書く。」

p.96 l.7 「 T^2 は F_n^1 に従うことを意味する。」 「 T^2 は F_n^1 に従うことを意味する。なお、本章では F_n^m や T^n を分布の記号として用いる。」

p.96 表 • 「 $\alpha = 5\%$ 」 「 $\alpha = 0.05$ 」
• 「 $\alpha = 1\%$ 」 「 $\alpha = 0.01$ 」

p.97 l.5 「(表の第1桁)」 「(表の第1列)」

p.97 l.-3 「正規母集」と「団」の間の改行 トル

pp.95-113 (例としてあげた各都市の気温のデータを第2版第1刷では <http://www.nakamuu.jp/kion/kion.htm> から引用したが、気象庁 web <http://www.jma.go.jp/jma/index.html> が更新されたのに気づいてそちらからデータを取り寄せたところ、いくつか違っていた。どちらのデータが正しいか(あるいは両方間違いか)検証するすべは持たないが、第2刷では公的に責任があるはずの気象庁のデータに基づくことに方針を変更した。このため pp.95-113 の数値のいくつかが変わる。ただし大きな変更はなく、本書の論旨は変わらない。また、例題の数値という意味では元の数値のままでも例題として問題はない。)

(1) p.95 表 「(旧)」

	2003年			2002年		
	仙台	東京	名古屋	仙台	東京	名古屋
7月26日	17.8	26.7	29.8	32.3	33.9	33.8
7月27日	22.4	27.3	29.5	30.6	34.5	34.9
7月28日	19.3	25.7	26.4	26.4	29.6	33.5
7月29日	22.4	25.4	27.2	26.9	30.9	32.9
7月30日	22.5	27.1	27.0	32.7	33.2	35.5
7月31日	26.0	30.9	31.8	32.1	34.8	36.8
8月1日	24.3	28.4	31.8	36.1	35.6	37.1
8月2日	26.6	31.4	32.9	31.7	35.2	35.0
8月3日	29.4	32.0	32.2	23.5	32.3	36.1
8月4日	29.7	33.4	34.6	26.1	31.4	35.9
8月5日	27.9	33.3	35.0	30.3	34.1	36.4
8月6日	24.9	30.9	34.1	35.2	35.7	38.2
8月7日	24.9	31.3	32.5	32.6	34.3	37.6
8月8日	26.3	31.0	32.0	33.6	35.7	35.7
データサイズ	14	14	14	14	14	14
標本平均	24.6	29.63	31.2	30.72	33.66	35.67
不偏分散	12.11	7.66	7.91	13.62	3.72	2.40

」 「(新)」

	2003 年			2002 年		
	仙台	東京	名古屋	仙台	東京	名古屋
7月26日	17.8	26.7	29.8	32.3	33.9	33.8
7月27日	22.4	27.3	29.5	30.6	34.5	34.9
7月28日	19.3	25.7	26.4	26.4	29.6	33.5
7月29日	22.4	26.4	27.5	26.9	31.0	32.9
7月30日	23.9	27.1	27.0	32.7	33.2	35.5
7月31日	26.0	30.9	31.8	32.1	34.8	36.8
8月1日	24.3	28.4	31.8	36.1	35.6	37.1
8月2日	26.6	31.4	32.9	31.7	35.2	35.0
8月3日	29.4	32.0	32.2	23.5	32.3	36.1
8月4日	29.7	33.4	34.6	26.1	31.4	35.9
8月5日	28.6	33.3	35.0	30.3	34.1	36.4
8月6日	25.2	31.0	34.1	35.2	35.7	38.2
8月7日	24.9	31.3	32.5	32.6	34.3	37.6
8月8日	26.3	31.0	32.0	33.6	35.7	35.7
データサイズ	14	14	14	14	14	14
標本平均	24.77	29.71	31.22	30.72	33.66	35.67
不偏分散	12.19	7.10	7.74	13.62	3.68	2.40

」

(引用を気象庁のものに変えたところ違いが多発)

(2) p.99 l.-12 「不偏分散が $V_X(\omega) = 12.11, V_Y(\omega) = 7.66, V_Z(\omega) = 7.91$ 」 「不偏分散が $V_X(\omega) = 12.19,$

$V_Y(\omega) = 7.10, V_Z(\omega) = 7.74$

(3) p.99 l.-6 「 $\frac{V_X}{V_Y}(\omega) = \frac{12.11}{7.66} = 1.58$ を比べて 」 「 $\frac{V_X}{V_Y}(\omega) = \frac{12.19}{7.10} = 1.72$ を比べて 」

(4) p.99 l.-5- -1 「1.58」 「1.72」 (4カ所)

(5) p.100 l.9 「1.58」 「1.72」

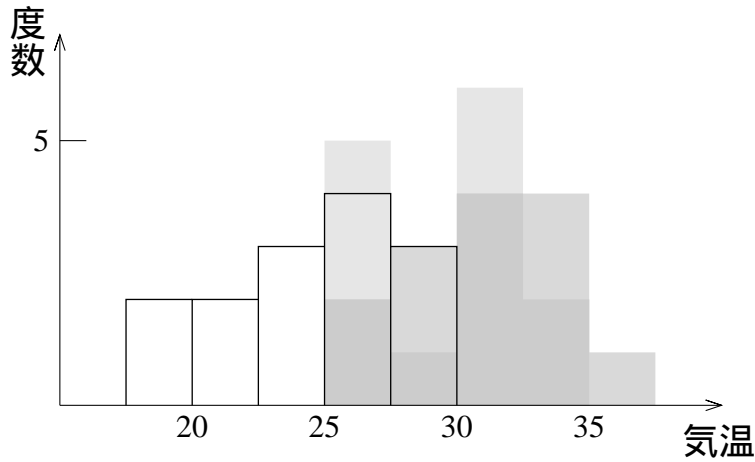
(6) p.100 l.13-15 「仙台と東京の標本平均の差は $\bar{X}_n(\omega) - \bar{Y}_n(\omega) = -5.0$. $V_X(\omega) = 12.11, V_Y(\omega) = 7.66, m = n = 14$ とともに定理 23 を適用すると, $0.8415(\mu_Y - \mu_X - 5.0)$ が T_{26} に従うことがわかる 」

「仙台と名古屋の標本平均の差は $\bar{X}_n(\omega) - \bar{Z}_n(\omega) = -6.5$. $V_X(\omega) = 12.19, V_Z(\omega) = 7.74, m = n = 14$ とともに定理 23 を適用すると, $0.8381(\mu_Y - \mu_X - 6.5)$ が T_{26} に従うことがわかる 」

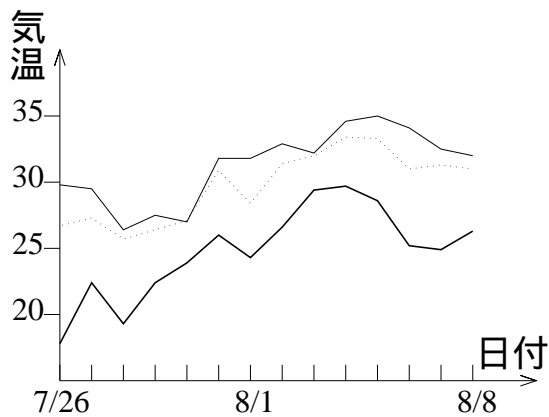
(7) p.100 l.18-21 「2003 年について (仙台 - 東京) の平均気温差の推定値は, -5.0 ± 2.4 (95% CL), -5.0 ± 3.3 (99% CL), となる . 同様に (仙台 - 名古屋) は -6.6 ± 2.5 (95% CL), -6.6 ± 3.3 (99% CL) , (東京 - 名古屋) は -1.6 ± 2.2 (95% CL), -1.6 ± 2.9 (99% CL) , を得る 」 「2003 年について (仙台 - 名古屋) の平均気温差の推定値は -6.5 ± 2.5 (95% CL), -6.5 ± 3.3 (99% CL) , となる . 同様に (仙台 - 東京) は -4.9 ± 2.4 (95% CL), -4.9 ± 3.3 (99% CL) (東京 - 名古屋) は -1.5 ± 2.1 (95% CL), -1.5 ± 2.9 (99% CL) , を得る 」

(平均は微妙に変わったが区間幅はこの精度では不変, ただし表題に合わせて仙台 - 名古屋を先にした)

(8) p.101 グラフ 下記で置き換える .



(9) p.102 グラフ 下記で置き換える .



(10) p.105 ℓ.16-20 「 $\bar{X}_{\text{仙台}}^B(\omega) - \bar{X}_{\text{名古屋}}^B(\omega) = -6.6, x^2 = 2 \times 16.25 \times 32.5$, を得るので信頼区間は -6.6 ± 1.3 (95% CL), -6.6 ± 1.8 (99% CL), となって, §4. よりも範囲が左右1度以上狭まって強い主張になっている. 同様に, (仙台 - 東京) は -5.0 ± 0.9 (95% CL), -5.0 ± 1.3 (99% CL), (東京 - 名古屋) は -1.6 ± 0.6 (95% CL), -1.6 ± 0.9 (99% CL) .」 「 $\bar{X}_{\text{仙台}}^B(\omega) - \bar{X}_{\text{名古屋}}^B(\omega) = -6.5, x^2 = 2 \times 17.46 \times 34.93$, を得るので信頼区間は -6.5 ± 1.3 (95% CL), -6.5 ± 1.9 (99% CL), となって, §4. よりも範囲が左右1度以上狭まって強い主張になっている. 同様に, (仙台 - 東京) は -4.9 ± 0.9 (95% CL), -4.9 ± 1.3 (99% CL) (東京 - 名古屋) は -1.5 ± 0.6 (95% CL), -1.5 ± 0.9 (99% CL) .」

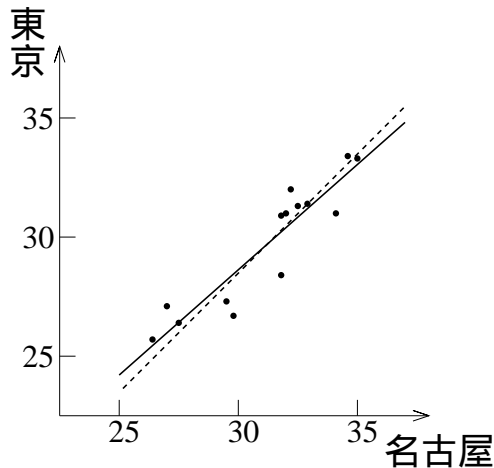
(11) p.106 表 「(旧) 計算すると」 「計算すると」

2003年	仙台	名古屋	\bar{X}_i^A
7月26日	17.8	29.8	23.8
7月27日	22.4	29.5	26.0
7月28日	19.3	26.4	22.9
7月29日	22.4	27.2	24.8
7月30日	22.5	27.0	24.8
7月31日	26.0	31.8	28.9
8月1日	24.3	31.8	28.1
8月2日	26.6	32.9	29.8
8月3日	29.4	32.2	30.8
8月4日	29.7	34.6	32.2
8月5日	27.9	35.0	31.5
8月6日	24.9	34.1	29.5
8月7日	24.9	32.5	28.7
8月8日	26.3	32.0	29.2
\bar{X}_j^B	24.6	31.2	$\bar{X} = 27.9$

(12) p.111 ℓ.5 「 $\hat{a} = 0.910, \hat{b} = 1.25$ を得る .」 「 $\hat{a} = 0.884, \hat{b} = 2.11$ を得る .」

2003年	仙台	名古屋	\bar{X}_i^A
7月26日	17.8	29.8	23.8
7月27日	22.4	29.5	26.0
7月28日	19.3	26.4	22.9
7月29日	22.4	27.5	25.0
7月30日	23.9	27.0	25.5
7月31日	26.0	31.8	28.9
8月1日	24.3	31.8	28.1
8月2日	26.6	32.9	29.8
8月3日	29.4	32.2	30.8
8月4日	29.7	34.6	32.2
8月5日	28.6	35.0	31.8
8月6日	25.2	34.1	29.7
8月7日	24.9	32.5	28.7
8月8日	26.3	32.0	29.2
\bar{X}_j^B	24.8	31.2	$\bar{X} = 28.0$

(13) p.111 グラフ 下記で置き換える .



(14) p.112 l.-9 「 $\hat{a} = 0.910$ を得ていた」 「 $\hat{a} = 0.884$ を得ていた」

(15) p.112 l.-7 「 $\frac{-0.090 \times 10.14}{\sqrt{14.45/12}} = -0.070$ が T_{12} に従うはずである」 「 $\frac{-0.116 \times 10.03}{\sqrt{13.74/12}} = -0.091$ が T_{12} に従うはずである」

(16) p.112 l.-4 「最適な実線 (傾き 0.91)」 「最適な実線 (傾き 0.884)」

(17) p.112 l.-3 「0.070 という数字は非常に」 「0.091 という数字はたいへん」

(18) p.113 l.-8 「 $S_{tot} = 99.59$ と $S_r = 85.14$ を得るので $R = 0.925$ となる」 「 $S_{tot} = 92.31$ と $S_r = 78.57$ を得るので $R = 0.923$ となる」

p.108 l.7 「実験などのデータ解析では背後にある法則やモデルを見つけることが目標である」 「実験などのデータ解析では背後にある法則やモデルを見つけること (モデリング) が目標である」

p.109 l.10 「これもいつものことだが $Z_{x,i}, Z_{y,i}$ たちは独立な確率変数とする」 の次に以下を挿入:

「以下簡単のため, 独立変数 x_i は測定誤差を伴わないとして, $Z_{x,i} = 0$ とおく」

($Z_x = 0$ で始めないと $Z = Z_y - aZ_x$ の a が処理できないので, 最初から入力変数は $Z_x = 0$ とすべきだった誤りの訂正.)

p.109 l.12 「 $E[Z_{x,i}] = E[Z_{y,i}] = 0, i = 1, \dots, n,$ 」 「 $E[Z_{y,i}] = 0, i = 1, \dots, n,$ 」

p.110 l.1 「 $Z_i = Z_{y,i} - aZ_{x,i}$ とおくと」 「 $Z_i = Z_{y,i}$ とおくと」

p.111 l.-3 「簡単のため」 「引き続き簡単のため」

p.112 l.-12 「この定理の証明は章末の補足に回す。」 「 $\hat{a} - a^*$ を複雑な式で割る理由は、ノイズ (測定誤差) の大きさ v が不明なので分布が v によらない量が必要だからである。この定理の証明は章末の補足に回す。

p.112 l.-12 - -11 「回帰値の検定ができる。」 「回帰値の検定や区間推定ができる。」

p.112 l.-3 - -1 「もっとも、夏の2週間のデータでだけでは温度の広がり狭すぎで優劣をつけられないのが真相だから、検定らしくするには範囲を広げて調べるべきである。」 「天気が大きなスケールで決まる夏のデータだから一見自然な結論に見えるかもしれないが、実はこの年は最高気温が低い日があって相関が強調されているという特殊事情による。」

(計算は合っていたが、結果の解釈を誤っていたので訂正。教科書では明示されることが少ないが、統計学の結論と称してとんでもない主張をする危険の実例である。統計学を実地に適用するときこのようなことを気をつけなければいけない。)

p.113 l.-6, -5, -1 • 「 $r(x, y)$ 」 「 \tilde{R} 」 (2カ所)

• 「 r 」 「 \tilde{R} 」

(第3章の相関係数 $r(x, y)$ は母分布統計量、第9章の標本相関係数は確率変数のサンプルだから同じ記号は紛らわしいから変える。)

p.113 l.-2 「対応物」 「対応物 (推定量)」

p.121 問1 「(9.13) を証明せよ。」 「決定係数が標本相関係数に等しいこと (9.13) を証明せよ。」

p.121 問2 「 2^* 」 「2」

p.122 l.3-11 「第8章までは主に1変数の正規母集団について ... 実用上も重要である。」 「本書の前半では標本平均 \bar{X}_n と不偏分散 V_n を用いて母平均と母分散の推定・検定を行う方法を紹介してきた。すでに紹介した推定や検定の概念そのものは一般性があるが、平均や分散以外の母集団特性値を母数とする検定や推定の一般論も必要である。本章ではそのような一般論の代表として尤度を用いる統計的推測の方法を紹介する。」

(本章で扱う定理は $n \rightarrow \infty$ の極限をとる漸近理論だから、動機付けは小標本理論ではなく大標本理論。)

p.125 l.-13 「 $I(Q|P) = \int_{\mathbb{R}^N} \left(\log \frac{f_Q(x)}{f_P(x)} \right) f_Q(x) dx$ 」

「 $I(Q|P) = \int_{\mathbb{R}^N} \frac{f_Q(x)}{f_P(x)} \left(\log \frac{f_Q(x)}{f_P(x)} \right) f_P(x) dx$ 」

(元の式でも同値だが、 f_P が基準であることがはっきりする式に改訂する。)

p.125 l.-12 「 Q を基準にして P が」 「 P を基準にして Q が」

p.126 l.3 「 θ に関する」 「 h に関する」

p.134 l.1 「偏差に直せば」 「偏差 \sqrt{v} は」

p.159 l.1-3 「このモデルでは適正な価格 (証券会社もこの人もぼろ儲けしない価格) が決まる (価格が決まる理由となる2項モデルの性質を2項モデルの完備性という。)」 トル

(説明を分かりやすくするために後ろに回す。)

p.159 l.-4 - p.160 l.1 「満期時までの成り行きによって値が変わる商品に現在時刻で値段が ... 複製 (複製ポートフォリオ) とは、このオプションを市場にある安全債券と株券の組み合わせで表すことをいう。」 「オプションは満期時までの成り行きによって価値が変わるけれども、市場にある原資

産（安全債券と1社の株券）の組み合わせでその価値を表す（複製する，複製ポートフォリオを組む）ことで以下のように値段が決まる．この事実を2項モデルにおける市場の完備性という．以下で決まる値段が2項モデルにおけるこのオプションの証券会社にとっての原価である（現実においてこの原価に比べてやたら高い価格や手数料で商品を勧められたら金融機関の努力不足である．）（改段落）

p.160 l.14 「となる。」と「以後。」の間に以下を挿入：

「オプション1単位が安全債券 C_B 枚と株券 C_S 枚の合計に等価なので価格も等しいとしてオプション価格を決めた．等価ならば等価格という性質が守られている市場は裁定(裁定取引, arbitrage)が無い，無裁定であると言う（改段落）」

p.162 l.-4 「 $C_S(k)B(k)$ 」 「 $C_S(k)S(k)$ 」

p.179 l.17 「のもとで注意した」 「の下で注意した」

p.191 l.6 「のもとに注意した」 「の下に注意した」

p.192 l.-6 「のもとに潜り込む」 「の下に潜り込む」

p.209 第9章 問1の略解 「 $r(x, y)^2$ を得る。」 「 \tilde{R}^2 を得る。」

p.209 第9章 問1と問3の略解の間 問2の略解を挿入する：

「回帰係数は $\hat{a} = 0.766$ および $\hat{b} = 6.35$ ．決定係数は $S_{tot} = 47.85$ と $S_r = 18.27$ から $R = 0.618$ ． $S_e = S_{tot} - S_r = 29.58$ と $\sqrt{D} = 5.58$ および $n - 2 = 12$ を用いれば帰無仮説 $a^* = 1$ の t 検定ができる． -0.069 が T_{12} に従うべきだが，本文と同様に棄却できないから $a^* = 1$ はデータと矛盾しない．全国的な異常低温があった2003年に比べて2002年は平年どおり高温で，直線関係で説明できる部分が少ないため， $\hat{a} = 0.766$ は1からやや大きくずれているけれども傾き1の直線がそれに比べて当てはまりが悪いとは言えない（つまり，傾きについてははっきりした傾向は無い）ことを意味する．」

p.212 第13章 問1の略解 「(13.11)のもとで注意したとおり」 「(13.11)の下で注意したとおり」

pp.215–218 参考文献を完全に入れ替える．以下を参照．

参考文献

[1] 服部哲也，理工系の確率・統計入門，学術図書，2005．

[2] 松本裕行，宮原孝夫，数理統計入門，学術図書，1990．

統計学の入門的教科書は恐ろしくなるほど多い．本書と同じ出版社にも [1, 2] などがある．特に [1] の著者は本書著者と同姓で読みは同名の別人（いきさつは [28] を参照）．以下，紙数の（すなわち，値段の）都合もあってバランスある紹介の余裕が無いので，増刷や改訂の際に少しずつリストを更新することをお許しいただきたい．

たとえば [3, 4, 5, 6, 7, 8, 9] が入門教科書や少し進んだ（大学学部水準くらいの）教科書のごくごく一部である．[3] は高校程度の数学知識だけを仮定していねいに数学的にギャップなしに解説した教科書．本書で省略したスターリングの公式の初等的な証明も含まれている．問題にはいねいな解答付き．[4] は具体例から入ることで引き込まれる構成になっている．連載終了後に教科書として出版されると聞いている．[5, 6] は演習書，[9] は詳しい（結果として分厚く，したがって，高い）入門的教科書．[10] は統計学そのものについて初等的に指摘した示唆的な薄い本．絶版は残念．本書第11章はこの本に触発された．[11] はよい教科書らしいが，未入手のまま絶版なのは残念．[12] は専門家による統計学の歴史書．数式を使わず，背後の精神に踏み込みながら，最新の話までカバーしている．[13] は高エネルギー実験データを集計した報告集．本書第6章のシミュレーションは，この報告の最初にある年次変化のグラフを理解することが著者の動機にあった．

- [3] 小針峴宏, 確率・統計入門, 岩波書店, 1973.
- [4] 渡辺浩, 使うための確率論入門, サイエンス社「数理科学」, 2006年6月号から連載.
- [5] 高岡浩一郎, 藤田岳彦, 穴埋め式確率・統計らくらくワークブック, 講談社サイエンティフィク, 2003.
- [6] 国沢清典編, 確率統計演習 (2分冊) 1 確率, 2 統計, 培風館, 1966.
- [7] 岡部靖憲, 確率・統計, 朝倉書店, 2002.
- [8] 長尾壽夫, 栗木進二, 数理統計学, 共立出版, 2006.
- [9] R. V. Hogg, A. T. Craig, J. McKean, *Introduction to Mathematical Statistics*, Prentice Hall, 2004.
数理統計学ハンドブック, 豊田秀樹 監訳, 朝倉書店, 2006.
- [10] 楠岡成雄, 確率・統計, 森北出版 (新数学入門シリーズ7), 1995.
- [11] 宮沢光一, 経済分析と決定理論, 東洋経済新報社 (経済分析のための数学入門5), 1971.
- [12] D. サルツブルグ, 統計学を拓いた異才たち, 竹内恵行, 熊谷悦生共訳, 日本経済新聞社, 2006.
- [13] Particle Data Group, *Review of particle properties*, Review of Modern Physics, **48–2**, Part II (1976) S1–S246; Rosenfeld, A.H., 1975, Ann. Rev. Nucl. Sci. 25, 555.

数学に詳しいほうが数理統計学の理解も明快で早くなる。たとえば [14] は確率論を本格的に学ぶための入門教科書。これらの教科書を含め、20世紀の確率論は測度論 (ルベグ積分論) に基づいて組み立てられることで数学的にすっきりし強力な解析手段となった。ルベグ積分の本格的入門教科書としてたとえば [15] を薦める。

本書第12,13章は時間の関数の集合上の確率論, つまり確率過程論が背景にある。確率過程論に基づく数理ファイナンスの教科書としてたとえば [16] がある。ランダムウォークはブラウン運動に比べると数学の初心者にはわかりやすいが, 数学的美しさには欠ける。藤田先生はこの数学者心理の隙間について, ブラウン運動の高度な数学的理論を片端からランダムウォークに美しく翻訳することに成功した [17]。ランダムウォークは通常は増分の独立同分布性が解析の要だが, [18] は自己相似性と指数を中心においたくりこみ群解析を解説している。

- [14] 熊谷隆, 確率論, 共立出版 (新しい解析学の流れ), 2003.
- [15] 吉田伸生, ルベグ積分入門 - 使うための理論と演習, 遊星社, 2006.
- [16] 関根順, 数理ファイナンス, 培風館, 確率論教程シリーズ7, 2007.
- [17] 藤田岳彦, ランダム・ウォーク, 日本評論社, 数学セミナー, 2004.4–2005.3, 2005.7–2005.12 (全18回連載)。この研究に関する本格的な教科書は: T. Fujita, M. Yor, *A discrete introduction to stochastic calculus and its applications to mathematical finance*, 近刊。
- [18] 服部哲弥, ランダムウォークとくりこみ群 — 確率論から数理物理学へ —, 共立出版 (新しい解析学の流れ), 2004.

(擬似)乱数生成アルゴリズムは原理・実用両面から深い研究が行なわれてきた。入門書としてたとえば [19] がある。[20] は, 本書原稿も利用する組版ソフト $\text{T}_{\text{E}}\text{X}$ の発明者としても有名な計算機科学の大家による伝説的名著。一様乱数について, 線形合同法は直前の数値から次の数値を定義したが, 直前の複数列のうまい線形結合を考えた方が性能が向上する (M 系列乱数)。周期の長い M 系列乱数としてメルセンヌ・ツイスター MT19937 などが発見されている [21]。一様乱数が生成できれば原理的にはそれを利用して種々の分布に従う乱数が生成できるが, 効率的に行なうには分布ごとに工夫を要する。[22] はそのための辞書本。絶版は残念。[23] は乱数とは何かという原理的問題に確率論の立場から一つの説得力ある解答を与え, その立場に即してモンテカルロ数値積分に向けた乱数を与えた。

- [19] 伏見正則, 確率的方法とシミュレーション, 岩波書店 (岩波講座応用数学 方法 10), 1994.
- [20] D. E. Knuth, *The art of computer programming*, 2nd ed., vol. 2, Addison-Wesley, Reading, MA, 1981.
- [21] M. Matsumoto, T. Nishimura, *Mersenne twister: A 623 dimensionally equidistributed uniform pseudorandom number generator*, ACM transactions on modeling and and computer simulations **8** (1998) 3–30.
- [22] L. Devroye, *Non-Uniform Random Variate Generation*, Springer, New York, 1986.
- [23] H. Sugita, *Robust numerical integration and pairwise independent random variables*, Jour. Comput. Appl. Math., **139** (2002) 1–8.

キーワードから web を検索すれば種々の比較的新しい統計データの手軽な検索ができる。本書で引用した web サイトは [24, 25, 26]。[24] から統計調査一覧のページに入れば厚生労働省統計表データベースがある。気温などの気象データは [25] から気象統計情報の中の過去の気象データ検索に入る。地震関係の資料は [26] から引用した。都道府県警の捜査費は MSN 毎日インタラクティブの記事 (2006 年 5 月 5 日 3 時 00 分更新, 毎日新聞) の引用だったがリンクが切れた。記事表題「都道府県警捜査費: 3 分の 1 に激減」で検索していただきたい。アクチュアリー資格試験問題は解答付きの試験問題集をアクチュアリー会が販売している [27]。

- [24] 厚生労働省, 厚生労働省統計表データベース
<http://www.dbtk.mhlw.go.jp/toukei/index.html>
- [25] 気象庁 <http://www.jma.go.jp/jma/index.html>
- [26] 地震調査研究推進本部, 宮城県沖地震の長期評価
<http://www.jishin.go.jp/main/chousa/00nov4/miyagi.htm>
 同, 長期的な地震発生確率の評価手法について, 2001
<http://www.jishin.go.jp/main/choukihyoka/01a/chouki0103.pdf>
- [27] 日本アクチュアリー会 <http://www.actuaries.jp/>
- [28] は本書の web ページ。訂正や関連資料や講義の様子などの情報を置いてある。
- [28] 服部哲弥, 「統計と確率の基礎」
<http://www.math.tohoku.ac.jp/~hattori/gakjutu.htm>