

正誤表

『基礎からのベイズ統計学—ハミルトニアンモンテカルロ法による実践的入門—』（豊田秀樹編著／朝倉書店，2015）の正誤表です。

[2016.1.15 公開， 5.10 最終更新]

1. 第3刷で修正された箇所

- p.vii, L.13
[誤] 波平釣果問題 ... [p.61] / [正] 波平釣果問題 ... [p.64]
- p.5, L.5
[誤] ... 異論のないの議論は ... / [正] ... 異論のない議論は ...
- p.35, L.5, 式 (2.34)
[誤] $F(x|\alpha, \beta) = \frac{1}{\beta - \alpha}, \quad 0 \leq x \leq 1$ /
[正] $F(x|\alpha, \beta) = \frac{1}{\beta - \alpha}, \quad \alpha \leq x \leq \beta$
- p.51, L.4
[誤] $V[X] = 0.016 (= 0.8 \times 0.2) \dots$ / ~~[正] $V[X] = 0.16 (= 0.8 \times 0.2)$~~
[再訂正] $V[X] = 0.016 (= 0.8 \times 0.2/10)$
- p.60, L.17-18
[誤] ... 1000 件回って, ... せいぜい 3 軒くらい ... /
[正] ... 1000 軒回って, ... せいぜい 3 件くらい ...
- p.65, L.9
[誤] ... 自然共役事前分がガンマ分布であること ... /
[正] ... 自然共役事前分布がガンマ分布であること ...
- p.68, 脚注 16 L.1
[誤] ... ベイズ定理の正規化係数 (3.4) 式 /
[正] ... ベイズ定理の正規化定数 (3.4) 式
- p.73, 問 11 L.2
[誤] ... 1 時間に 4 人のポアソン分布に ... /
[正] ... 1 時間に 4 人のガンマ分布に ...

- p.87, 最終行
[誤] ... 事後分布の正規化係数 / [正] ... 事後分布の正規化定数
- p.127, 最終行から次ページ L.1
[誤] 効果量は平均値の差が本当に意味のあるものかどうかを判断する指標です。 /
[正] 効果量は, 平均値間の差や, 目標値と平均値との差を標準偏差の単位で評価する指標です。
- p.133, L.23
[誤] ... 2 つの群で母平均が 1 秒以上差が開く確率 /
[正] ... 2 つの群の母平均の差が 1 秒以上となる確率
- p.200
[以下を追記]

■ 参考文献 (B.1, B.2)

- Gelman, A. & Rubin, D. B. (1992): Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, **7**, 457–511.
- Gelman, A. (1996): Inference and monitoring convergence. In Gilks, R. W., Richardson, S. & Spiegelhalter, J. D. (Eds.), *Markov Chain Monte Carlo in Practice*, Chapman & Hall. 131–143.
- Geweke, J. (1992): Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid & A. F. Smith (Eds.), *Bayesian Statistics*, **4**, Oxford University Press. 169–193.
- Stan Development Team. (2014): Stan Modeling Language Users Guide and Reference Manual, Version 2.5.0.
- 伊庭幸人, 種村正美, 大森裕浩, 和合肇, 佐藤整尚, 高橋明彦 (2005): 『計算統計 II —マルコフ連鎖モンテカルロ法とその周辺—』, 岩波書店。

2. 第4刷で修正された箇所

- p.130, 表 6.3
データ no.30 → [誤] 内径 45.46 / [正] 内径 145.46
- p.134, L.6
[誤] ... 平均値の差が 0 以上であるという ... /
[正] ... 平均値の差が 0 より大きいという ...

- p.134, L.12
[誤] ... 一定の値 c 以上であるという ... /
[正] ... 一定の値 c より大きいという ...
- p.135, L.11-12
[誤] ... 実験群よりも 1 秒以上遅くなる ... /
[正] ... 実験群と比較して 1 秒より大きい ...
- p.137, L.24
[誤] ... 研修前よりも 2 件以上多い ... /
[正] ... 研修前と比較して 2 件より多い ...
- p.141, L.7
[誤] kg 以上であるという研究仮説と, 2 kg 以上であるという ... /
[正] kg より大きいという研究仮説と, 2 kg より大きいという ...
- p.145, 最終 L.3 から次ページ L.2
[誤] したがって, この生成量 $\delta^{(t)}$ の事後分布の 95%確信区間が 0 を含んでいなければ, $(\lambda_B - \lambda_A) > 0$ といってよいでしょう. $\delta = (\lambda_B - \lambda_A)$ の 95%確信区間を見ると, ... が分かります. この区間が 0 を含んでいないことから $\lambda_B > \lambda_A$ であると主張してもよさそうです. /
[正] [1 行削除] $\delta = (\lambda_B - \lambda_A)$ の 95%確信区間を見ると, ... が分かります. ただし, この場合 $\delta^{(t)}$ については上側に関心がないたため, 事後分布の下側 2.5%点 (95%下側) が 0 より大きければ, 97.5%の確信で $(\lambda_B - \lambda_A) > 0$ と言えます. 今回は下側 2.5%点が $0.11 > 0$ なので $\lambda_B > \lambda_A$ であると主張してもよさそうです.
- p.164, L.10
[誤] 差の確信区間は 0 を含んでいます. /
[正] 興味の対象である下側 2.5%点は 0 より小さいため, 97.5%の確信で差があるとは結論づけられませんでした.
- p.186, 左 L.13
[誤] 次に, 2つの母数の差を表す生成量 δ の 95%確信区間は $[-0.415, 0.842]$ であり, 0 をまたいでいます. つまり, 2つの母数の差は, 95%の確率で -0.415 から 0.842 の間に存在するというを意味し, $\lambda_B - \lambda_A < 0$ にもなり得る

と解釈できます. /

[正] 次に, 2つの母数の差を表す生成量 δ の 95%確信区間 $[-0.415, 0.842]$ より, 下側 2.5%点は -0.415 で 0 より小さいので, 97.5%の確信では差があるとは言えません.

- p.206, プログラム解説 L.9-10
[誤] ... `data` にデータファイルを指定します. /
[正] ... `data` に 4 行目で作成したリストオブジェクトを指定します.

3. 第 5 刷で修正された箇所

- p.93, L.2, 8, 11, 13 の 4 箇所および p.99 演習問題 7
[誤] `sd` / [正] 分散
- p.115, L.13
[誤] L と ϵ を両方とも大きくすると /
[正] L を大きく, ϵ を小さくすると
- p.119, 最終 L.3
[誤] $p^{(2)} = -2.35$ / [正] $p^{(3)} = -2.35$
- p.146, L.5 (7.4) 式 第 3 項条件部分
[誤] $\delta > 0$ / [正] $\delta^{(t)} > 0$
- p.156, 最終 L.12 (7.31) 式 第 3 項条件部分
[誤] $\zeta_{0.70} < 450$ / [正] $\zeta_{0.70}^{(t)} < 450$
- p.156, 最終 L.4
[誤] $u_{\zeta_{0.70} < 450}^{(t)}$ / [正] $U_{\zeta_{0.70} < 450}$
- p.163, 最終 L.3

$$[誤] q = \frac{1}{1 - \rho^2} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right) - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right) \right]$$

$$[正] q = \frac{1}{1 - \rho^2} \left[\left(\frac{x_1 - \mu_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1 - \mu_1}{\sigma_1} \right) \left(\frac{x_2 - \mu_2}{\sigma_2} \right) + \left(\frac{x_2 - \mu_2}{\sigma_2} \right)^2 \right]$$

- p.173,L.5 (8.17) 式 第2項
[誤] $g(\sigma_{\alpha r}^2(t), \sigma_{\beta r}^2(t), \sigma_{er}^2(t))$ / [正] $g(\sigma_{\alpha r}^{2(t)}, \sigma_{\beta r}^{2(t)}, \sigma_{er}^{2(t)})$
- p.183, 右段, 付録1 章末問題 第4章 図直下
[誤] 9) 問題9でシミュレートした... /
[正] 9) 省略
10) 問題9でシミュレートした...
- p.184, 付録1 章末問題 第5章 6) D
[誤] $D: v(\tau) = \sqrt{\frac{2K}{m}} = \sqrt{\frac{3920}{9.8}} = 20$ /
[正] $D: v(\tau) = \sqrt{\frac{2K}{m}} = \sqrt{\frac{3920}{100}} \simeq 6.26$
- p.184, 付録1 章末問題 第5章 10)
[誤] シミュレートをした結果 /
[正] 初期値 $\theta^{(1)} = 0.5$ としてシミュレートをした結果

4. 7.3節の計算上の不備に関して (5刷以降修正済)

☆ 7.3節に計算上の不備が判明しました (4刷以前)。読者にお詫びするとともに関連する 7.4, 7.5節を含む pp.148-153, ならびに章末問題の解答 pp.186-188 を以下のように訂正し, 該当する青字箇所を修正した版面 (pp148-153.pdf, pp186-188.pdf) を別途公開いたします。[朝倉書店編集部]

- p.148, 最終 L.9-最終行
[誤] $\alpha = 3$ のガンマ分布を考えます。... 悪ければ1時間以上待つ必要があることが分かりました。 /
[正] $\alpha = 3, \lambda = \lambda^{(t)}$ のガンマ分布から, T 個の予測値を以下のように生成します。(7.11) 式中の \sim は, 左辺の変数が右辺の確率分布に従うということを表す記号です。

$$x^{*(t)} \sim f(3, \lambda^{(t)}) \quad (7.11)$$

ただし, 個々の $x^{*(t)}$ はガンマ分布に従いますが, T 個の $x^{*(t)}$ によって構成される分布はガンマ分布とはならず, 新たに観測されるデータ x^* の予測分

布の近似となります。この予測値に, 単位時間である5をかけた $x^{*(t)} \times 5$ について T 個の平均値を求めることで, 流れ星を3つ観測するまでの平均的な待ち時間を得ることができます。

実際に推定した結果が表 7.4 です。 $x^* \times 5$ の推定結果から, H君がこれから3個流れ星を観測するには, 平均して約30分待つ必要があり, 少なくとも5分程度待たなくてはならず, 高々1時間半程度待てばよいことがわかりました。

- p.149, 表 7.4

[誤]

	EAP	post.sd	95%下側	95%上側
η	29.932	14.589	12.886	67.562

[正]

	平均	標準偏差	95%下側	95%上側
$x^* \times 5$	29.903	24.258	4.693	92.413

- p.150, 表 7.5 L.4

[誤] $U_{p_{x=3} > 1/36}$ 0.871 0.335 0.000 1.000 /
[正] $U_{\theta > 1/36}$ 0.896 0.305 0.000 1.000

- p.151, L.7-L.8

[誤] 3本買って, ちょうど3本目で当たりが出る確率 /
[正] 当たりが出る確率 θ

- p.151, L.9

[誤] $u_{p_{x=3} > 1/36}^{(t)}$ / [正] $u_{\theta > 1/36}^{(t)}$

- p.151, L.10 (7.17) 式

[誤] $u_{p_{x=3} > 1/36}^{(t)} = g(\theta^{(t)}) = \begin{cases} 1 & p_{x=3} > 1/36 \\ 0 & \text{それ以外の場合} \end{cases}$

[正] $u_{\theta > 1/36}^{(t)} = g(\theta^{(t)}) = \begin{cases} 1 & \theta^{(t)} > 1/36 \\ 0 & \text{それ以外の場合} \end{cases}$

- p.151, L.11

[誤] $u_{p_{x=3}>1/36}^{(t)}$ は, (7.16) 式で定めた $p_{x=3}^{(t)}$ /

[正] $u_{\theta>1/36}^{(t)}$ は, $\theta^{(t)}$

- p.151, L.12

[誤] $u_{p_{x=3}>1/36}^{(t)}$ / [正] $u_{\theta>1/36}^{(t)}$

- p.151, L.12

[誤] $p_{x=3} > 1/36$ / [正] $\theta > 1/36$

- p.151, L.13

[誤] ちょうど3本目で当たりが出る確率 /

[正] Tくんのデータから推測される当たりが出る確率

- p.151, L.14

[誤] $U_{p_{x=3}>1/36}$ / [正] $U_{\theta>1/36}$

- p.151, L.15

[誤] ちょうど3本目で当たりが出る確率 /

[正] 当たりが出る確率 θ

- p.151, L.16

[誤] 87.1% / [正] 89.6%

- p.152, L.3(7.19) 式 第1項

[誤] $f(x|\theta)$ / [正] $f(x|k, \theta)$

- p.153, 表 7.6

[誤]

	EAP	post.sd	95% 下側	95% 上側
θ	0.273	0.092	0.113	0.471
p	0.202	0.103	0.043	0.434
ν	8.369	3.640	4.247	17.623

[正]

	EAP	post.sd	95% 下側	95% 上側
θ	0.273	0.092	0.113	0.471
p	0.202	0.103	0.043	0.434
	平均	標準偏差	95% 下側	95% 上側
$x^* + 2$	8.349	6.754	2.000	26.000

- p.153, L.6

[誤] 分布関数 / [正] 確率関数

- p.153, L.10-L.12

[誤] また, 95%確信区間は [0.043, 0.434] です. 5通中2通以上合格する確率は平均的には2割程度と推測されましたが, その確信区間が広いため, 運が悪ければ約4%の確率でしか2通以上合格しない, ということです. /

[正] 【上記記述を削除】

- p.153, 最終 L.9-最終行

[誤] この問2に答えるためには, ... 最大18通送ると良いでしょう. /

[正] 問2に答えるためには, 予測分布を利用します. $k=2, \theta=\theta^{(t)}$ の負の2項分布に従う予測値

$$x^{*(t)} \sim f(2, \theta^{(t)}) \quad (7.23)$$

を生成し, この予測値に成功回数である2を足した $x^{*(t)} + 2$ によって, 2回成功するまでの総試行数に関する予測分布を近似的に得ることができます.

表(7.6)で, 総試行数を表す $x^* + 2$ の平均値は8.349となっており, 平均的には9通エントリーシートを送ればそのうち2通合格することが期待できます. しかし, できるだけ確実に2社の合格を得るためには, 予測分布の区間 [2.000, 26.000] の95%上側を参照し, 26通送ると良いでしょう.

- p.186, 付録1 章末問題 第7章 2) 本文 L.4-5

[誤] 最短で約90m, 最長で約300m 歩けば次の八重桜を見ることができます. /

[正] 少なくとも約90m 歩かないと次の八重桜を見ることができませんが, 高々約300m 歩けば次の八重桜を見ることができます.

- p.186, 付録1 章末問題 第7章4) 本文 L.3-最終行

[誤] ν の EAP 推定値は 6.820, 確信区間は (3.600, 14.102) となったため, 平均的には約 7 回挑戦すれば新たに 2 つぬいぐるみを取ることができ, 多く見積もって 15 回の挑戦で確実に目的を果たせる可能性が高いことが分かりました. /

[正] $x^* + 2$ の平均値は 6.772, 予測分布の区間は [2.000, 21.000] となったため, 平均的には約 7 回挑戦すれば新たに 2 つぬいぐるみを取ることができ, 余裕をもって 21 回挑戦すれば目的を果たせる可能性が高いことが分かりました.

- p.187, 付録1 章末問題 第7章4) 表

[誤]

	EAP	post.sd
p	0.104	0.084
ν	6.820	2.838
	95% 下側	95% 上側
p	0.008	0.324
ν	3.600	14.102

[正]

	EAP	post.sd
p	0.104	0.084
	95% 下側	95% 上側
p	0.008	0.324
	平均	標準偏差
$x^* + 2$	6.772	5.319
	95% 下側	95% 上側
$x^* + 2$	2.000	21.000

- p.220, 付録4 Stan コード 右段 L.6-7

[誤] `real<lower=0.000001> eta;`
`eta <- 5*3/lambda;` /

[正] `real<lower=0.000001> pred;`
`pred <- gamma_rng(3,lambda)*5;`

- p.220, 付録4 Stan コード 右段最終 L.11

[誤] `P <- if_else(p3>0.028,1,0);` /
 [正] `P <- if_else(theta>0.028,1,0);`

- p.221, 付録4 Stan コード 左段 L.6

[誤] `real nu;` /
 [正] `real beta;`
`int pred;`

- p.221, 付録4 Stan コード 左段 L.11

[誤] `nu <- 2*(1-theta)/theta+2;` /
 [正] `beta <- theta/(1-theta);`
`pred <- neg_binomial_rng(2, beta)+2;`

5. 第6刷で修正予定の箇所

- p.51, L.4

[誤] $V[X] = 0.16 (= 0.8 \times 0.2)$ / [正] $V[X] = 0.016 (= 0.8 \times 0.2/10)$

- p.137, (6.10) 式

[誤] $f(x_1, x_2 | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho)} e^{-q/2}$ /

[正] $f(x_1, x_2 | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-q/2}$

- p.163, 8.2.1 項 L.4

[誤] $f(x_1, x_2 | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho)} e^{-q/2}$ /

[正] $f(x_1, x_2 | \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-q/2}$