

# Errata

(Mathematical Introduction to Data Science by Sven A. Wegner)

October 15, 2024

- Page 8, Line -4:

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^n x_i y_i - \bar{y} \sum_{i=1}^n x_i - \bar{x} \sum_{i=1}^n y_i + n \bar{x} \bar{y}$$

- Page 9, Line -3:

$$0 = \sum_{i=1}^n (ax_i + b - y_i) = a \sum_{i=1}^n x_i + nb - \sum_{i=1}^n y_i = an\bar{x} + n\mathbf{b} - n\bar{y},$$

- Page 13, Line 10: ... constant random variable **av**.

- Page 15, Line -9:  $\dots + \frac{2}{n^2} \sum_{i < j} x_i x_j \mathbb{E}(\mathcal{E}_i) \mathbb{E} \mathcal{E}_j)$

- Page 15, Line -5: since  $\overline{x^{(n)^2}} = \mathbf{var}(x^{(n)}) + \overline{x^{(n)}}^2$  as the sum ...

- Page 14, Line -2: If the latter is the case, then  $\text{sign}(r_{xy}) = \text{sign}(\langle \mathbf{u}, \mathbf{v} \rangle) = \dots$

- Page 27, Line 22: ..., the rounded ~~logistie~~ **logistic** regressor, ...

- Page 39, Line 10: *The calculation of the  $k$ -nearest neighbors of  $x$  can be implemented such that at most  $(n \cdot d \cdot \mathbf{k})$ -many mulitplikations have to be carried out.*

- Page 39, Line 14: In the Euklidean metric, it requires  $(d - 1)$ -many multiplications to compute one distance if we omit the root, which we can do as it does not change the argmin. This leads to

$$(d - 1) \cdot (n + (n - 1) + \dots + (n - k + 1)) \leq \mathbf{C} \cdot d \cdot k \cdot n$$

multiplikationen ~~with a suitable~~  **$C \in \mathbb{N}$** .

- Page 40, Line 3: ..., we choose  $k$ -nearest neighbors  $x_1, \dots, x_{\mathbf{k}}$  of  $x$  and denote their labels by  $y_1, \dots, y_{\mathbf{k}}$ .

- Page 40, Line 9:  $f: X \rightarrow Y, f(x) = \frac{\sum_{i=1}^k w(x_i, x) \cdot y_i}{\sum_{i=1}^k w(x_i, x)}$

- Page 41, Line 5:  $\tilde{x}^{(i)} = \left( \mathbf{a} + \frac{(x_1^{(i)} - \min_{j=1, \dots, n} x_1^{(j)})(b - a)}{\max_{j=1, \dots, n} x_1^{(j)} - \min_{j=1, \dots, n} x_1^{(j)}}, \dots \right)$

- Page 41, Line 7:  $\tilde{x}^{(i)} = \left( \frac{x_1^{(i)} - \overline{x_1^{(\mathbf{c})}}}{\sigma_1}, \dots \right)$

- Page 42, Line 14:  $\rho(x^{(1)}, x^{(4)}) = \mathbf{3.681}$

- Page 46, Line 1: We discuss some of these methods in Exercise **3.10**.

- Page 45, Line 28: We thus see that texts no. 1 and text no. 2 are significantly more cosine similar than text no. 1 and text no. **3** or text no. **2** and text no. 3.

- Page 48, Line 26: The cosine distance, on the other hand, **may appear hear more natural, as the scalar product increases if the frequency of the fixed word increases in the second text.**

- Page 52, Line 11: For **finite subsets**  $A, B \subseteq X$  we define ...

- Page 53, Line 20:

```

1: function LINKAGE-BASED CLUSTERING ( $X, \rho, D, \delta$ )
2:    $k \leftarrow \#D$ 
3:   for  $i \leftarrow 1$  to  $k$  do
4:      $C_i \leftarrow \{x_i\}$ 
5:   while  $\min_{i \neq j} \rho(C_i, C_j) \leq \delta$  and  $k \geq 2$  do
6:      $m \leftarrow 0$ 
7:      $(i^*, j^*) \leftarrow \operatorname{argmin}_{i \neq j} \rho(C_i, C_j)$ 
8:     for  $\ell \leftarrow 1$  to  $k - 1$  do
9:       if  $\ell = \min(i^*, j^*)$  then
10:         $C_\ell \leftarrow C_{i^*} \cup C_{j^*}$ 
11:       if  $\ell = \max(i^*, j^*)$  then
12:         $m \leftarrow 1$ 
13:         $C_\ell \leftarrow C_{\ell+m}$ 
14:       else
15:         $C_\ell \leftarrow C_{\ell+m}$ 
16:      $k \leftarrow k - 1$ 
17:   return  $C_1, \dots, C_k$ 

```

- Page 54, Line -6:  $K: \mathcal{C}_k \rightarrow \mathbb{R}$
- Page 56, Line -16: The following pseudocode *approximates* a minimizer of the  $k$ -means cost function.
- Page 56, Pseudocode:

```

1: function K-MEANS ( $D, k, X, \rho$ )
2:    $\mu_1, \dots, \mu_k \leftarrow$  pairwise different points from  $X$ 
3:   for  $i \leftarrow 1$  to  $k$  do
4:      $C_i \leftarrow \{x \in D \mid i \in \operatorname{argmin}_{j=1, \dots, k} \rho(x, \mu_j)\}$ 
5:    $U \leftarrow \text{True}$ 
6:   while  $U = \text{True}$  do
7:      $U \leftarrow \text{False}$ 
8:     for  $i \leftarrow 1$  to  $k$  do
9:        $\mu_i \leftarrow \mu(C_i)$ 
10:    for  $i \leftarrow 1$  to  $k$  do
11:       $C'_i \leftarrow \{x \in D \mid i \in \operatorname{argmin}_{j=1, \dots, k} \rho(x, \mu_j)\}$ 
12:      if  $C'_i \neq C_i$ 
13:         $C_i \leftarrow C'_i$ 
14:         $U \leftarrow \text{True}$ 
15:   return  $C_1, \dots, C_k$ 

```

In the lines 4 and 9 of the pseudocode we pick as a single  $i$  in the case that the argmin is not unique.

- Page 57, Line 7:

$$\mu(A) \in \operatorname{argmin}_{\mu \in \mathcal{A}} \sum_{x \in A} \rho(x, \mu)^2, \quad \text{respectively} \quad \mu(A) \in \operatorname{argmin}_{\mu \in \mathcal{X}} \sum_{x \in A} \rho(x, \mu).$$

- Page 57, Line -7: *Proof.* For  $j \geq 1$  denote by  $(C_1^{(j)}, \dots, C_k^{(j)})$  that clustering which the algorithm produces in the  $j$ -th round.

- Page 57, Line -4:  $K(C_1^{(j)}, \dots, C_k^{(j)}) = \min_{\mu_1, \dots, \mu_k \in \mathcal{X}} \sum_{i=1}^k \sum_{x \in C_i^{(j)}} \rho(x, \mu_i)^2$

- Page 58, Line 3: ...line 11 of Algorithm 4.9 ...

- Page 58, Line 7: There, Picture 1b corresponds to the penultimate line in the estimate and Picture 2a to the line above that.

- Page 58, Line 10: ...as we have just moved point  $x_3$  from cluster  $C_2$  in Figure 1b to cluster  $C_1$  in Figure 2a, ...

- Page 59, Line -4: We assume that we start with the **initial values**  $\mu_1 = 2$  and ...
- Page 63, Line -2: ... In Example 5.7,  $\lambda_2 \neq 0$  and there are no clusters (**or, depending on how one prefers to see it, one single cluster**), in Example 5.8 ...
- Page 67, Line 8: **For the other direction let  $\{v_1, \dots, v_n\}$  be a basis consisting of eigenvectors corresponding to the  $\lambda_i$  and let  $U \subseteq \mathbb{R}^n$  be a subspace with  $\dim U = n - k + 1$ . By construction  $U \cap \text{span}\{v_1, \dots, v_k\} \neq \{0\}$  and we can select  $0 \neq x = \alpha_1 v_1 + \dots + \alpha_k v_k \in U$ . Then it follows**

$$\frac{\langle x, Mx \rangle}{\langle x, x \rangle} = \frac{\sum_{i=1}^k \lambda_i \alpha_i^2}{\sum_{i=1}^k \alpha_i^2} \leq \frac{\lambda_k \sum_{i=1}^k \alpha_i^2}{\sum_{i=1}^k \alpha_i^2} = \lambda_k,$$

since the  $\lambda_i$ 's are increasing. **With this we get  $\min_{0 \neq x \in U} \frac{\langle x, Mx \rangle}{\langle x, x \rangle} \leq \lambda_k$  which then leads to**

$$\max_{\substack{U \subseteq \mathbb{R}^n \\ \dim U = n-k+1}} \min_{\substack{x \in U \\ x \neq 0}} \frac{\langle x, Mx \rangle}{\langle x, x \rangle} \leq \lambda_k.$$

- Page 68, Line 14: Let  $G = (V, E)$  be a graph with  $\deg(v) > 0$  for all  $v \in V$ .
- Page 70, Line 10:

$$\lambda_2(\mathcal{L}) = \min_{\substack{x \neq 0 \\ \langle Dx, \mathbf{1} \rangle = 0}} \frac{\sum_{\{i,j\} \in E} (x_i - x_j)^2}{\sum_{i=1}^n x_i^2 d_i}.$$

- Page 70, Line 16:

$$\mathcal{L}D^{1/2}\mathbf{1} = D^{-1/2}LD^{-1/2}D^{1/2}\mathbf{1} = D^{-1/2}L\mathbf{1} \stackrel{\substack{\uparrow \\ \text{Proposition} \\ 5.6}}{=} D^{-1/2}0\mathbf{1} = 0D^{1/2}\mathbf{1}.$$

- Page 70, Line 18:

$$\begin{aligned} \lambda_2(\mathcal{L}) &\stackrel{\text{Theorem 5.13}}{=} \min_{\substack{x \neq 0 \\ \langle x, D^{1/2}\mathbf{1} \rangle = 0}} \frac{\langle x, \mathcal{L}x \rangle}{\langle x, x \rangle} \stackrel{(*)}{=} \min_{\substack{y \neq 0 \\ \langle D^{1/2}y, D^{1/2}\mathbf{1} \rangle = 0}} \frac{\langle D^{1/2}y, \mathcal{L}D^{1/2}y \rangle}{\langle D^{1/2}y, D^{1/2}y \rangle} \\ &= \min_{\substack{y \neq 0 \\ \langle y, D\mathbf{1} \rangle = 0}} \frac{\langle y, Ly \rangle}{\langle y, Dy \rangle} = \min_{\substack{y \neq 0 \\ \langle Dy, \mathbf{1} \rangle = 0}} \frac{\sum_{\{i,j\} \in E} (y_i - y_j)^2}{\sum_{i=1}^n y_i^2 d_i} \end{aligned}$$

- Page 71, Line 14: *(ii)  $\min(\text{vol } S_k, \text{vol } S_k^c) = \text{vol } S_k^c$  and  $\text{vol } S_k^c - \text{vol } S_{k+1}^c = d_{k+1}$  hold whenever  $r \leq k \leq n-1$ .*
- Page 71, Line 17:

$$\text{vol } S_k^c - \text{vol } S_{k+1}^c = \sum_{i=k+1}^n d_i - \sum_{i=k+2}^n d_i = d_{k+1}$$

- Page 72, Line 9:

$$\langle Dx, \mathbf{1} \rangle = \sum_{i=1}^n d_i x_i = \dots$$

- Page 72, Line -1 (and Page 73, Line 1):

$$\begin{aligned} \dots &= \text{vol } S - \text{vol } S \cdot \frac{\text{vol } S}{\text{vol } S + \text{vol } S^c} \\ &\geq \text{vol } S - \text{vol } S \cdot \frac{\text{vol } S}{2 \text{vol } S}, \end{aligned}$$

- Page 73, Line 9: ... and our goal in the following will be to show  $\lambda_2 \geq \alpha^2/2$ , ...
- Page 73, Line 12 (Equation (5.2)):

$$\dots \text{ and } \langle Dx, \mathbf{1} \rangle = \sum_{i=1}^n d_i x_i = 0$$

- Page 73, Line 21:

$$\begin{bmatrix} x_1 - x_r \\ \vdots \\ x_{r-1} - x_r \\ 0 \\ x_{r+1} - x_r \\ \vdots \\ x_n - x_r \end{bmatrix} = \begin{bmatrix} x_1 - x_r \\ \vdots \\ x_{r-1} - x_r \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ x_r - x_{r+1} \\ \vdots \\ x_r - x_n \end{bmatrix} =: p - n.$$

- Page 74, Line –3 (until top of page 75):

$$\begin{aligned} \lambda_2 &= \frac{\sum_{\{i,j\} \in E} (x_i - x_j)^2}{\sum_{i=1}^n x_i^2 d_i} \\ &\geq \frac{\sum_{\{i,j\} \in E} ((p_i - p_j)^2 + (n_i - n_j)^2)}{\sum_{i=1}^n (p_i^2 + n_i^2) d_i} \\ &\stackrel{(5.3)}{\geq} \frac{\sum_{\{i,j\} \in E} (p_i - p_j)^2 + \sum_{\{i,j\} \in E} (n_i - n_j)^2}{\sum_{i=1}^n p_i^2 d_i + \sum_{i=1}^n n_i^2 d_i} \\ &\stackrel{(5.4)}{=} \frac{\sum_{\{i,j\} \in E} (p_i - p_j)^2 + \sum_{\{i,j\} \in E} (n_i - n_j)^2}{\sum_{i=1}^n p_i^2 d_i + \sum_{i=1}^n n_i^2 d_i} \\ &\geq \min \left( \frac{\sum_{\{i,j\} \in E} (p_i - p_j)^2}{\sum_{i=1}^n p_i^2 d_i}, \frac{\sum_{\{i,j\} \in E} (n_i - n_j)^2}{\sum_{i=1}^n n_i^2 d_i} \right) \\ &\stackrel{(5.5)}{\geq} \min \left( \frac{\sum_{\{i,j\} \in E} (p_i - p_j)^2}{\sum_{i=1}^n p_i^2 d_i}, \frac{\sum_{\{i,j\} \in E} (p_i + p_j)^2}{\sum_{\{i,j\} \in E} (p_i + p_j)^2}, \dots \right) \\ &=: \min \left( \frac{Z}{N}, \dots \right). \end{aligned}$$

- Page 75, Line 8:

$$N = \sum_{i=1}^n p_i^2 d_i \cdot \sum_{\{i,j\} \in E} (p_i + p_j)^2 \geq \sum_{i=1}^n p_i^2 d_i \cdot \sum_{\{i,j\} \in E} 2(p_i^2 + p_j^2)$$

- Page 76, Line 3:

$$\dots = \left( \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbb{1}_E(i, j) \sum_{k=i}^{j-1} (p_k^2 - p_{k+1}^2) \right)^2$$

$\uparrow$   
 telescopic  
 sum

- Page 79, Line 13: Finally, we want to note that Theorem 5.21 together with Remark 5.17(i) provides an upper bound for the eigenvalue  $\lambda_2(\mathcal{L})$ .

- Page 99, Line –3: ~~Multiplication with  $V^T$  from the left in~~

**Corollary 5.22.** Let  $G = (V, E)$  be a graph with  $\deg(i) > 0$  for all  $i \in V$ . Then for the second smallest eigenvalue  $\lambda_2$  of the normalized Laplace matrix of  $G$  the estimate  $\lambda_2 \leq 2$  holds.  $\square$

- Page 105, Zeile –1 and Page 106, Line 1:

$$\begin{aligned} A &= \underbrace{\begin{bmatrix} 0.07 & 0.29 & 0.32 & 0.51 & 0.66 & 0.18 & -0.23 \\ 0.13 & -0.02 & -0.01 & -0.79 & 0.59 & -0.02 & -0.06 \\ 0.68 & -0.11 & -0.05 & -0.05 & -0.24 & 0.56 & -0.35 \\ 0.15 & 0.59 & 0.65 & -0.25 & -0.33 & -0.09 & 0.11 \\ 0.41 & -0.07 & -0.03 & 0.10 & -0.02 & -0.78 & -0.43 \\ 0.07 & 0.73 & -0.67 & 0.00 & -0.00 & 0.00 & 0.00 \\ 0.55 & -0.09 & -0.04 & 0.17 & 0.17 & -0.11 & 0.78 \end{bmatrix}}_U \underbrace{\begin{bmatrix} 12.4 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 9.5 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.3 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}}_\Sigma \underbrace{\begin{bmatrix} 0.56 & 0.09 & 0.56 & 0.09 & 0.59 \\ -0.12 & 0.69 & -0.12 & 0.69 & 0.02 \\ -0.40 & -0.09 & -0.40 & -0.09 & 0.80 \\ -0.41 & -0.09 & -0.40 & -0.09 & 0.80 \\ 0.51 & 0.48 & -0.51 & -0.48 & -0.00 \\ 0.48 & -0.51 & -0.48 & 0.51 & -0.00 \end{bmatrix}}_{V^T} \\ &= \begin{bmatrix} 0.07 & 0.29 & 0.32 \\ 0.13 & -0.02 & -0.01 \\ 0.68 & -0.11 & -0.05 \\ 0.15 & 0.59 & 0.65 \\ 0.41 & -0.07 & -0.03 \\ 0.07 & 0.73 & -0.67 \\ 0.55 & -0.09 & -0.04 \end{bmatrix} \begin{bmatrix} 12.4 & & \\ & 9.5 & \\ & & 1.3 \end{bmatrix} \begin{bmatrix} 0.56 & 0.09 & 0.56 & 0.09 & 0.59 \\ -0.12 & 0.69 & -0.12 & 0.69 & 0.02 \\ -0.40 & -0.09 & -0.40 & -0.09 & 0.80 \end{bmatrix}. \end{aligned}$$

- Page 106, Zeile 6:

$$\check{A} = \begin{bmatrix} 0.15 & 1.97 & 0.15 & 1.97 & 0.56 \\ 0.92 & 0.01 & 0.92 & 0.01 & \textcolor{red}{0.94} \\ 4.84 & 0.03 & 4.84 & 0.03 & 4.95 \\ 0.36 & 4.03 & 0.36 & 4.03 & 1.20 \\ 2.92 & \textcolor{red}{-0.00} & 2.92 & \textcolor{red}{-0.00} & 2.98 \\ -0.34 & 4.86 & -0.34 & 4.86 & 0.65 \\ \textcolor{red}{3.92} & \textcolor{red}{0.02} & \textcolor{red}{3.92} & \textcolor{red}{0.02} & \textcolor{red}{4.00} \end{bmatrix} = \begin{bmatrix} 0.07 & 0.29 \\ 0.13 & -0.02 \\ 0.68 & -0.11 \\ 0.15 & 0.59 \\ 0.41 & -0.07 \\ 0.07 & 0.73 \\ 0.55 & \textcolor{red}{-0.09} \end{bmatrix} \begin{bmatrix} 12.4 & & & & \\ & 9.5 & & & \end{bmatrix} \begin{bmatrix} 0.56 & 0.09 & 0.56 & 0.09 & 0.59 \\ -0.12 & 0.69 & -0.12 & 0.69 & 0.02 \end{bmatrix}$$

- Page 107, Line 15:

$$= \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} 0.07 & 0.29 & \dots & -0.23 \\ 0.13 & -0.02 & \dots & -0.06 \\ 0.68 & -0.11 & \dots & -0.35 \\ 0.15 & 0.59 & \dots & 0.11 \\ 0.41 & -0.07 & \dots & -0.43 \\ 0.07 & 0.73 & \dots & 0.00 \\ 0.55 & \textcolor{red}{-0.09} & \dots & 0.78 \end{bmatrix} \begin{bmatrix} 12.4 & & & & \\ & 9.5 & & & \\ & & 1.3 & & \\ & & & \ddots & \end{bmatrix} \begin{bmatrix} 0.56 & 0.09 & 0.56 & 0.09 & 0.59 \\ -0.12 & 0.69 & -0.12 & 0.69 & 0.02 \\ \vdots & & & & \vdots \\ 0.48 & -0.51 & -0.48 & 0.51 & 0.00 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

- Page 109, Line -2:

$$u_2 = 0.29 \cdot \text{Abbie} - 0.02 \cdot \text{Bailey} + \dots - \textcolor{red}{0.09} \cdot \text{Gladys},$$

- Page 109, Line 10:

$$\check{A} = \begin{bmatrix} 0.07 & 0.29 \\ 0.13 & -0.02 \\ 0.68 & -0.11 \\ 0.15 & 0.59 \\ 0.41 & -0.07 \\ 0.07 & 0.73 \\ 0.55 & \textcolor{red}{-0.09} \end{bmatrix} \begin{bmatrix} 12.4 & & & & \\ & 9.5 & & & \end{bmatrix} \begin{bmatrix} 0.56 & 0.09 & 0.56 & 0.09 & 0.59 \\ -0.12 & 0.69 & -0.12 & 0.69 & 0.02 \end{bmatrix}$$

Abbie →
Alien
Casablanca
Star Wars
Titanic
The Matrix

Bailey →
Catherine
Darlene
Elena
Fatima
Gladys

- Page 229, Line 13: Alternatively, with **sigmoid** activation, ...

- Page 282, Line 6: (ii) For  $A, B \in \Sigma$  with  $\mathcal{P}(B) \neq 0$ ,  $\mathcal{P}(A|B) := \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)}$  ...

- Page 283, Line 19:

$$\rho(x) = \frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\frac{\|x-\mu\|^2}{2\sigma^2}} \quad \text{respectively} \quad \rho(x) = \frac{1}{\lambda^{\textcolor{red}{d}}(B)} \cdot \mathbf{1}_B(x),$$

- Page 286, Line -5: For  $A = A_1 \times \dots \times A_d \subseteq \mathbb{R}^d$  with  $A_i \in \mathcal{B}^{\textcolor{red}{d}}$  we calculate

- Page 288, Line 2:

$$\begin{aligned} (\rho_1 * \rho_2)(s) &= \frac{1}{2\pi\sqrt{ab}} \int_{\mathbb{R}} \exp\left(-\frac{(s-t)^2}{2a}\right) \exp\left(-\frac{t^2}{2b}\right) dt \\ &= \frac{1}{2\pi\sqrt{ab}} \int_{\mathbb{R}} \exp\left(-\frac{b(s^2 - 2st + t^2) + at^2}{2ab}\right) dt \\ &= \frac{1}{2\pi\sqrt{ab}} \int_{\mathbb{R}} \exp\left(-\frac{t^2(b+a) - 2stb + bs^2}{2ab}\right) dt \\ &= \frac{1}{\textcolor{red}{2\pi\sqrt{ab}}} \int_{\mathbb{R}} \exp\left(-\frac{t^2(b+a)/c - 2stb/c + bs^2/c}{2ab/c}\right) dt \\ &= \frac{1}{\sqrt{2\pi c} \sqrt{2\pi(ab/c)}} \int_{\mathbb{R}} \exp\left(-\frac{(t - (bs)/c)^2 - (sb/c)^2 + s^2(b/c)}{2ab/c}\right) dt \\ &= \frac{1}{\sqrt{2\pi c}} \exp\left(+\frac{(sb/c)^2 - s^2(b/c)}{2ab/c}\right) \frac{1}{\sqrt{2\pi(ab/c)}} \int_{\mathbb{R}} \exp\left(-\frac{(t - (bs)/c)^2}{2ab/c}\right) dt \\ &= \frac{1}{\sqrt{2\pi c}} \exp\left(+\frac{(sb/c)^2 c^2 - s^2(b/c)c^2}{2abc}\right) \\ &= \frac{1}{\sqrt{2\pi c}} \exp\left(+\frac{s^2(b^2 - bc)}{2abc}\right) \\ &= \frac{1}{\sqrt{2\pi c}} \exp\left(-\frac{s^2}{2c}\right), \end{aligned}$$