Bernhard Elsner – elsner@latroika.com

Steven J. Miller
Williams College
Williamstown, MA

13 septembre 2019

Concerns: Comments on your book *The Probability Lifesaver* (2017 print)

Dear Professor,

by chance I came across your excellent textbook in a Paris book store. I bought it because I saw immediately that it is written by someone who loves not only the subject but also the reader. Therefore I will recommend it for purchase to my University library. Certainly you are interested in some remarks in view of a 2nd edition. You will find them below.

p.10 The figure seems wrong.

p.21 In the Fibonacci example : Before using the Bring It Over Method one must first prove (or at least mention) that the series converges.

p.38 You don't mention that differentiating under the integral sign is not always allowed. Here you are less cautious than at other places where you warn about infinity, etc. ;-) By the way, the method works only in some very, very special and nice cases. How about

$$\int_0^2 \frac{x^5 - 1}{\log(x)} \,\mathrm{d}x \ ?$$

p.42 It's very good that you mention the origin of the notation $\mathbb{Z}$. But it is *Zahl* and not *zahl*, because in German language all nouns are written with an upper-case letter, without any exception. (By the way, in Danish it was the same until 1948.)

p.44 Here you use $\varsubsetneq$, but on page 50 you use $\subsetneq$.

p.51 Twice on this page you mention *circle* and *sphere*. Probably you mean *disk* and *(three dimensional) ball*.

p.56 One arrow seems to be missing in the figure.

p.56 I like the American way to write .7 for 0.7. But is there a rule ? Sometimes you drop the zero and some inches later you don't, all that in a single equation.

p.58 You ask the reader to prove $\Pr(\emptyset) = 0$, but you proved it already on page 54.

p.58 Should be $\Pr(\omega_k)$ instead of $\Pr(a_k)$.

p.59 Should be $A \subset \Omega$ instead of $A \in \Omega$.

p.66 dart at the unit circle $\rightharpoonup$ dart at the unit disk

p.70 I think it is misleading to say that the set $\mathcal{N}$ could look like $\{0, \sqrt{2}/2, \pi/4, \ldots\}$ because usually the three omission points are used for countable sets only.

p.73 It would be better to put the first phrase of exercise 2.9.6 into exercise 2.9.4.

p.75 Exercise 2.9.26 : Should be a sum, not an equation.

p.86 Should be $n!/(k!(n-k)!)$ and not $n!/k!(n-k)!$. See also page 202 where on two instances it is correct, and on one instance it is wrong. (Refer to my general remark about a/bc at the end of this list.)

p.90 $52!/5!47!$ $\rightharpoonup$ $52!/(5!47!)$. See also page 105.

p.90 Why begin 3.2.2 with the case of "Nothing" since it is more difficult ? And you promise to subtract the (small) errors later, but then you forget it.

p.91 You mention the "multiplication rule" but you have never stated it on the previous pages (or I missed it).

p.93 Here is another way to see why the answer 52% to problem 3.2.2 is wrong : On page 91 you showed that the disjoint event of having nothing special has probability of 50.7%.

p.96 We kept using binomial coefficients and not multiplying by numbers ; that's a good sign that we're respecting order. $\rightharpoonup$ We kept using binomial coefficients and not multiplying by numbers ; that's a good sign that we're *not* respecting order.

Sorry, no remarks on the following fifty pages, I skipped them ☺

p.160 Your definition of "circular orderings" isn't precise. When you write "relative orderings around a circular table" it means to me that two orderings are the same if everyone has the same two neighbours. So the result should be $(n-1)!/2$ (because of the mirror reflection interchanging left and right neighbours).

p.172 I only learned English at school (in Germany), so I don't understand why you write "their 13 cards" when it is just one person. Same with "their hand".

p.172 There's $4 \cdot 3$ ways or there are $4 \cdot 3$ ways ?

p.172 There is a whole paragraph in italics. You never did that on other pages.

p.172 all the probabilities below $\rightharpoonup$ all the cardinalities below

p.176 At the end of the first paragraph you say that our strategy is to calculate the non-derangements by calculating the derangements. But at the end of the fourth paragraph you propose the opposite strategy. That's a little bit confusing ! In fact, it's the second strategy you are pursuing.

p.176 You distinguish $\{2, 3, 4, 1\}$ and $\{3, 2, 4, 1\}$. But usually these are notations for the same set. It would be clearer to write the tuplets $(2, 3, 4, 1)$ and $(3, 2, 4, 1)$. Later, on page 180, you do that correctly.

p.184 There is a confusion with $n!$ and $!n$ in exercise 5.5.24.

p.188 When reading I didn't get the precise meaning of "have exactly four of their six digits equal" because I didn't understand that you wanted them also on the same positions. In the solution you say it better : "the plates match in exactly four places". (Moreover it would be clearer to begin with "One chooses randomly two car owners", because one might think that you ask for the probability that at least one pair of car owners amongst all have number plates which...)

p.195 "Solution to (1) with replacement" should be in italics like the other.

p.208 If I understand correctly the lottery system in the US sells all possible tickets, one for each. Here in Europe it's different, you choose the numbers you like. So there can be several winners who will then have to share the Jackpot. This is why spending your money on <u>all</u> possibilities is risky here.

p.208 Example 6.3.2 : You ask for the probability of winning, but in the solution you don't give it. Indeed, the results are not uniformly distributed.

p.210 You kindly ask the reader to spend a few minutes before reading on. So I did. Not finding any better idea, I followed the stupid but safe way and added the possibilities for all types of configurations (there are eleven types). That worked fine, here you can find the OpenOffice sheet with the formulae :

www.latroika.com/mathoman/uvsq/DenombrementBE03.ods

While doing so, I wondered. There is a much easier way to see things! What does a lottery result like $(1,1,1,6,8,8)$ could stand for? Well, it means that you gave three cookies to person number 1, one cookie to person number 6 and two cookies to person number 8. Thus all amounts to distributing 6 cookies to 50 persons and we get

$$\binom{6+50-1}{50-1} = \binom{55}{49} = \binom{55}{6}.$$

You see, you praise laziness — but I'm even lazier than you! I went on reading and understood your very witty but a little bit complicated reasoning. In fact, what you do is to prove again the symmetry relation $\binom{n}{n-k} = \binom{n}{k}$. It's a second proof by story of this relation; the first one (p.88) is a short story, but this one is a rather lengthy novel ☺

p.213 I fear that the partition by the sets $\mathcal{M}$ and $\mathcal{W}$ will soon be banned from math classes due to our colleagues from the department of gender studies.

p.227 The two figures don't belong to each other.

p.228 You mention a Theorem 2.7. Where is it?

p.230 "The argument above, etc." I don't see which argument you refer to.

p.238 At the end of chapter 7 and the beginning of chapter 8 you mention that it is easier to find a closed form for an integral than for a sum. Later on p.387 you admit that the tool kit from calculus is not enough to integrate any given expression.

pp.266, 267 It is confusing to have the tables not at the place where they are mentioned first in the text.

p.267 You glue two coins together and still call that tossing fair coins?

p.268 In the integral $\int_{-1}^{1} \frac{2}{\pi}\sqrt{1-x^2}$ the dx is missing. You calculate it by substitution ; but some lines earlier you solve the double integral geometrically (as area of the unit disk). So why not doing the same here and saying that the graph of the function $x \mapsto \sqrt{1-x^2}$ is the upper half of the unit circle ?

p.279 We've essentially prove $\rightarrow$ We've essentially proven

p.286 You say we can find the density of $Z = X + Y$ by evaluating the convolution integral. Why don't you do it ? For this example it is very easy (and instructive).

p.295 "nice interval I where X is defined." As X is defined on Ω you want to say "nice interval I containing all the values of X" (or simply $X(\Omega) \subset I$).

p.299 Same remark for "interval I where the random variable X is defined".

p.323 I don't see what is wrong in the purportedly wrong reasoning of exercise 11.7.2 (except that you write "for all x" at the end whereas it should have been at the beginning).
By the way it would be nicer if the counting of the exercises didn't include the section number (11.2 instead of 11.7.2). Hard to code since you still want the exercises to appear on the table of contents. But you seem to be a great coder... ☺

p.330 "One of the most powerful techniques in probability and combinatorics is to compute something two different ways." I would even go further ! The majority of proofs in all mathematics is based on that principle of doing the same thing two different ways. (As an illustration I often give the simple geometric proof that in a triangle the heights always meet in a point.)

p.330 Last equation : You forget to write the result for the sum, the variance $np(1-p)$.

p.331 You write $E(X)$ instead of $\mathbb{E}(X)$.

p.331 In this example the computation of the mean of X can be done faster : Let $Y \sim \mathcal{B}(4, \frac{1}{6})$. Then $P(X = -1) = P(Y = 0)$ and $\mathbb{P}(X = k) = \mathbb{P}(Y = k)$ for $k = 1, \ldots, 4$. Thus

$$\mathbb{E}(X) = \mathbb{E}(Y) - \mathbb{P}(X = -1) = 4 \times \frac{1}{6} - \frac{625}{1296} = \frac{239}{1296}.$$

p.334 $p_1^{x_1} p_2 x_2 \cdots p_k^{x_k}$ $\rightarrow$ $p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k}$

p.336 Your way of computing $\mathbb{P}(D \geqslant m)$ is far-fetched. If X follows a geometric distribution of parameter p then

$$\forall k \in \mathbb{N} : \qquad \mathbb{P}(X > k) = (1-p)^k,$$

since it's the probability that the first k outcomes are failures. No necessity to manipulate an infinite sum ☺

p.340 (p^k) $\rightarrow$ p^k

p.342 For an IKEA store your numbers of sold beds par day seem unrealistic to me (too small) !

p.374 In one of the integrals you have an exponential of u but you integrate by x.

p.374 $I(0,1) = \sqrt{2\pi}$ $\rightarrow$ $I(0,1) = 1$

p.374 In two of the integrals you have an exponential of x but you integrate by u.

p.387 The function $A(x)$ doesn't make the (best possible) point because it simplifies :

$$A(x) = \log\left(\cos\left(x^2+1\right)\exp\left(x+\sqrt{\sin(x)}\right)\right) = \log\left(\cos\left(x^2+1\right)\right) + x + \sqrt{\sin(x)}$$

Better omit the three letters exp in the expression of $A(x)$.

p.387 The whole discussion could be resumed in the following way : mathematicians give names to functions which they esteem important. For example, one of the anti-derivatives of $x \mapsto 1/\sqrt{1-x^2}$ is important, so we name it (and its name is arcsin because it's the inverse of another important function named sin).
Here is a question : I don't know (but maybe you do) what's is exactly going on inside a pocket calculator, when one asks it to find arcsin(0.4). Is it really helpful for the machine that arcsinus is a "named" (=known) function ? If not, it doesn't really make a difference to which functions we give names and to which we don't.

p.388 "It suffices to know how to compute $\int_0^x \phi(t)\mathrm{d}t$ for any fixed, positive x." Actually all would be clearer if you dropped the adjective "fixed". In fact, I misunderstood your phrase and thought you were saying "if for a fixed x_0 you know $\int_0^{x_0} \phi(t)\mathrm{d}t$ then you can deduce $\int_0^x \phi(t)\mathrm{d}t$ for any x."

p.391 There is a problem with the figure : a probability can't converge to 4.

p.394 In exercices 14 and 15 the Prob symbol shouldn't be in italics.

p.405 $n!/k!(n-k)! \rightarrow n!/(k!(n-k)!)$

p.406 e^{-x} times a polynomial in x $\rightarrow$ e^{-x} times a power function of x

p.408 In the figure there are five pairs of parameters but only four plots. Moreover it would be nice to know which corresponds to which.

p.412 15.7 Families of Random Variables $\rightarrow$ 15.7 Families of Distributions

p.414 I guess that a reader who was able to follow you until here never makes a mistake of the type shown in equation (15.1).

p.428 Same remark as p.408

p.432 The proof by grouping parentheses is known to all students that were able to follow you until here : it is called induction. ☺

p.438 It is confusing to call the unit sphere in $\mathbb{R}^n$ by the name "n-dimensional sphere". The dimension should be the one of the object, not of an (arbitrary) ambient space. So the sphere you mean is of dimension $n-1$, and the standard notation is S^{n-1}. But since this is not a geometry book, you don't need that notation. Just call it "the unit sphere in n-space".
At the bottom of the page, it's even more confusing : You mean "volume of a unit ball in n-space" (or "n-ball" if you prefer), but not "volume of a unit n-sphere" (because that would be an $(n-1)$-dimensional volume and would be zero when compared to the n-dimensional volume of the n-cube).
Maybe the best would be to define these objects once and for all on p.51 : the balls $B_a(r)$ and $\overline{B}_a(r)$ and the sphere $S_a(r) = \overline{B}_a(r) \setminus B_a(r)$.

p.441 $f_{\nu_1} * \cdots * f_{\nu_2}$ $\rightarrow$ $f_{\nu_1} * f_{\nu_2}$

p.451 X is at most a $\rightarrow$ X is at least a

p.452 to go below zero then the integrand can be negative $\rightarrow$ to go below zero, where the integrand is negative, and so the whole integral can become negative

p.454 Mistake in inequality : $> 1 - \frac{1}{k^2}$ $\rightarrow$ $\geqslant 1 - \frac{1}{k^2}$

p.454 at least 1 (at least 100%) $\rightarrow$ at most 1 (at most 100%)

p.459 The coefficient of $\mathrm{P}(A_1 \cap \cdots \cap A_n)$ ought be $(-1)^{n+1}$

p.452 At the beginning of the example, please state the conditions : there are 4 persons and 52 cards.

p.460 When you compute $\mathrm{P}(A_i \cap A_j)$ you make a counting error : you divide by $\binom{52}{26}$. But the correct number is $\binom{52}{13}\binom{39}{13}$. Or $\binom{52}{26}\binom{26}{13}$, if you prefer. Here is the good news : This correction makes the probability S_2 smaller and thus sharpens Bonferroni's inequality ☺

p.461 You write "F continuous". But isn't F even differentiable, like any cdf ?

p.462 In the first picture the labels of both axes aren't right. Should be 0.2, 0.4, etc. and 0.05, 0.10, etc.

p.463 "We denote this by" and you give twice the same.

p.469 Within two lines you write $-1/2$ once using the $ and the other time not using it : $-1/2$ vs. -1/2.

p.474 In the last line it should read "are all Bernoulli random variables".

p.475 Same remark for the beginning of the second paragraph.

p.475 and a -1 for tail (put the $ sign)

p.470 if we have k heads and $2N - k$ tails then S_{2N} equals $2k - 2N$.

p.477 The notation of the statement of the lemma is not rigorous. One should write $A(N) \overset{N\to\infty}{\sim} B(N)$; the asymptotic equivalence is declared as usual by :

$$f(x) \overset{x\to\infty}{\sim} g(x) \iff f(x) - g(x) \overset{x\to\infty}{=} \mathrm{o}(g(x)).$$

p.479 You write that $\log(x)$ is increasing for $x \geqslant 1$. In fact it is increasing for $x > 0$. Your inequality would be easier to understand if you took the sum of $n-1$ terms, i.e. $\sum_{k=2}^{n}$; in fact in both integrals you take $n-1$ intervals of length 1. The fact that you have n terms in your sum is confusing (only in this special case it doesn't matter because $\log(1) = 0$).

p.481 (1) we have an alternating series *whose terms are decreasing in absolute value* The fact (2) wasn't used before, you will use it just below.

p.481 "Now that you have seen the integration by parts attack". I didn't see anything of the kind !

p.484 In understand that $n = 2N$. But I don't see why

$$X = n^{n/2+n/4+n/8+\cdots+n/2^N} 2^{-n/2} 4^{-n/4} 8^{-n/8} \cdots (2^N)^{-n/2^N}$$

should be greater than

$$Y = n^{n/2+n/4+n/8+\cdots+n/2^N+n/2^N} 2^{-n/2} 4^{-n/4} 8^{-n/8} \cdots (2^N)^{-n/2^N} 2^{-n/2^N}.$$

We rather have the contrary since

$$\frac{Y}{X} = n^{n/2^N} 2^{-n/2^N} = n^1\, 2^{-1} = \frac{n}{2}.$$

p.490 Bottom of page : in the formula for $f(m)$ where is the parameter λ ?

p.494 As far as I remember my lectures in mechanics the generating functions in hamiltonian systems don't have much to do with the generating functions in combinatoris or probability treated in your book. What generating functions are you refering to when you write that they are used "throughout mathematical physics" ?

p.497 "We now use the quadratic formula." Why ? We don't look for s.

p.498 The last sum should be $\sum_{\ell=0}^{k}$.

p.502 Twice : $n!|s| > 1 \;\rightarrow\; n!|s|^n > 1$.

p.504 In the definition you say the same twice : X taking on values in the integers ; X is only non-zero at the integers.

p.505 important probability distributions (binomial, geometric, etc.)

p.513 Bottom of page : $M_X(t) = G_X(\log t) \;\rightarrow\; M_X(t) = G_X(e^t)$.

p.514 Bottom of page : for some positive δ

p.515 expected value of the product of etc.

p.517 for all s $\rightarrow$ for all $s \neq 0$

p.520 in Equation (19.2) $\rightarrow$ in equation (19.2)

p.428 In the last equation some n's are missing : f_n. You could also add : for $|t| < \delta$.

p.529 Just before 20.2 there is a 20.11.4 all alone.

p.530 Theorem : whose moment generating functions converge $\rightarrow$ whose moment generating function converges (there is only one, because identically distributed variables)

p.531 $(x_i - \mu)^2 \;\rightarrow\; (x_n - \mu)^2$

p.532 $f(x_i) \;\rightarrow\; f(x_n)$

p.540 we leave it to the reader to determine which approach they like more $\rightarrow$ we leave it to the reader to determine which approach (s)he likes more. Or : we leave it to the readers to determine which approach they like more. (? ? ?)

p.544 with mean 0 and variance 4 $\rightarrow$ with mean 3 and variance 4

p.557 You prove that $|e^{it}| = 1$. Do you seriously believe that someone who doesn't already know this basic fact understands one hundredth of what you write on those pages about the Fourier transform ?

p.558 What is I in the second integral on the page ?

p.558 can only be zero $\rightarrow$ can only be infinite

p.560 You use i as subscript of f_i. I recommend k because there are $2\pi i$ everywhere.

p.561 Same for X_i

p.562 $p(-x^2/2)/\sqrt{2\pi}$ $\rightarrow$ $\exp(-x^2/2)/\sqrt{2\pi}$

p.563 $\mathbb{E}[x^2]$ $\rightarrow$ $\mathbb{E}[X^2]$

p.574 You write $\mathbb{E}(X)$ and "the expected value of X", but the variable X was never introduced before

p.575 Since $\mathbb{V}(X) = \sigma^2/n$, to cut the variance in half you need to collect twice (not four times) as much data.

p.575 $\mathrm{N}(\mu, \sigma^2/n)$ $\rightarrow$ $N(\mu, \sigma^2/n)$ (take like italics on p.385)

p.576 This mean $\rightarrow$ This means

p.581 having x machines $\rightarrow$ having n machines

p.629 x_i as that varies with i $\rightarrow$ x_n as that varies with n
(I admit that I don't know why you mention this.)

p.581 $\displaystyle\sum_{m=1}^{\ell} \frac{1}{n}$ $\rightarrow$ $\displaystyle\sum_{m=1}^{\ell} \frac{1}{m}$; $1/2\ell$ $\rightarrow$ $1/(2\ell)$

p.683 The statements of your root and ratio tests are wrong. Why should there be a problem with a convergence radius equal to 1 ? Take the ratio test :

$$|s| < \rho = 1 \implies \lim_{n\to\infty} \left| \frac{a_{n+1}s^{n+1}}{a_n s^n} \right| = \rho|s| = |s| < 1 \implies \sum a_n s^n \text{ converges.}$$

So, instead of "if $\rho = 1$ then there is..." it should be "if $|s| = \rho$ then there is..."

p.684 small and diverges $\rightarrow$ smaller and diverges

p.692 zahl $\rightarrow$ Zahl (see p.42)

p.705 Twice the accent should go this way : L'Hôpital

p.708 In one of the quotients you write h instead of h_1.
ih_2/ih_2 $\rightarrow$ $ih_2/(ih_2)$
in the plane there's really only two ways to approach a point $\rightarrow$ on the line there are really only two ways to approach a point

p.709 Definition : if f has a series expansion $\rightarrow$ if in every point of U the function f has a series expansion. (Because U is not necessarily a disk.)

p.711 Twice : $1/\pi n$ $\rightarrow$ $1/(\pi n)$

General remark : All computer programs interpret `a/b*c` as

$$a/bc = \frac{a}{b}\,c \neq \frac{a}{bc}.$$

This is why I recommend $1/(\pi n)$ instead of your $1/\pi n$; and on p.269 simply write

$$2/\pi\sqrt{1-y^2} \quad \text{or} \quad \frac{2}{\pi}\sqrt{1-y^2}$$

instead of the superfluous parenthesis in

$$(2/\pi)\sqrt{1-y^2}\,.$$