

Errata and suggested changes for *Understanding Advanced Statistical Methods*, by Westfall and Henning

Special thanks to Robert Jordan and Arkapravo Sarkar for suggesting and/or catching many of these.

p. 1, 6 lines from bottom:

“... allow us to reverse-engineer natural processes ...” →

“ ... allow us to reverse-engineer the workings of natural processes ...”

p. 39, three lines from bottom:

... approximately normal distribution. →

... approximately a normal distribution.

p. 54, end of paragraph following the “Requirements for a Continuous pdf”:

... probability is 1.0% or 100%. →

... probability is 1.0 or 100%.

p. 57, Cell C2 Formula:

= EXP(-0.5\*(A2-74^2/18.7^2)) → = EXP(-0.5\*(A2-74)^2/18.7^2)

p. 85, first integral in the displayed equations:  $p(y)d \rightarrow p(y)dy$

p. 101, first sentence of last paragraph: ... Figure 4.4 ... → ... Figure 4.3 or 4.4 ...

p. 103, second para: ... set it off in bold so you can find it later: → ... set it off so you can find it later:

p. 107, first line:  $y_{(6)} = 2.3 \rightarrow y_{(5)} = 2.3$

p. 111, second line in Example 4.7: 0.00017076 → 0.00020019 ; 0.01121524 → 0.01118705

(Figures 4.11 and 4.12 are slightly affected by this change, but the difference is irrelevant to the point of assessing normality.)

p. 134, Exercise 5.4B, first line: ...simulate many 10 more ... → ...simulate 10 more ...

p. 153, first line below equation (6.10):

“By expressing the numerator with the reverse condition, using Equation 6.7, you get...” →

“Expressing the left hand side with the reverse condition and using Equation 6.7, you get...”

p. 157, three lines from bottom:

“The entire area of the entire rectangular region of Figure 6.8 is ...” →

“The entire area of the rectangular region of Figure 6.8 is ...”

p. 160, first sentence of third paragraph. Some might prefer the phrasing “...as more data come in” to the given phrasing “...as more data comes in.”

p. 172: Critique of the Population Interpretation of Probability Models → Critique of the Population Interpretation of Probability Models, Even When Data are Sampled From a Population

p. 172, delete “7.3.1 Even When Data Are Sampled from a Population”

p. 172 – 175: 7.3.2  $\rightarrow$  7.3.1; 7.3.3  $\rightarrow$  7.3.2; 7.3.4  $\rightarrow$  7.3.3; 7.3.5  $\rightarrow$  7.3.4; 7.3.6  $\rightarrow$  7.3.5

p. viii, in the Table of Contents: Make all the changes as indicated in the point immediately above.

p. 179, second row of Table 7.4:  $<1 \rightarrow \leq 1$ . Make the same change to Table 7.5 on p. 180 and Table 7.11 on P. 186.

p. 182, 3<sup>rd</sup> paragraph, second line: Before you collect your data ...  $\rightarrow$  Before you collect your data ...

p. 183, immediately before section 7.5, add the following paragraph.

Some sources call the process model an **infinite population model**. Using the infinite population model, many of the criticisms given previously in Section 7.3 are no longer relevant. However, it is not helpful to refer to infinite populations for several reasons: (i) The actual definition of the infinite population is ill-defined; e.g., what is the infinite population of coin tosses? How does it start, with heads or tails? Does it matter whether the coin is bent? Strike one against the infinite population model. (ii) The notion of data being somehow selected from some existing infinite population is pure fiction: Is there a set of items in existence from which the data are selected? The answer is almost always, “No.” For example, is tomorrow’s value of the Dow Jones Industrial Average somehow selected from some existing infinitely big bag of possible Dow Jones values? Of course not. Strike two against the infinite population model. (iii) The notion that there is an infinite population is consistent with the fact that probabilistic processes can produce an infinite amount of data. So this extra step of postulating an infinite population is simply unnecessary baggage. The postulation of a probabilistic process alone suffices. Strike three against the infinite population model. (iv) The notion that data are produced by a population is not scientific: Scientific enterprises postulate models for how data are produced. Processes define any “population” data, whether infinite or finite. So it is scientifically better to start with the notion of a data generating process, rather than a population, no matter whether finite or infinite. Strike four against the infinite population model. (v) If you want to demonstrate the meaning of an “infinite population” via simulation, you cannot do so directly by creating an infinite population and sampling from it. Instead, you simply have the computer produce data from some distribution, such as the normal distribution. But this is just the process model. Since the infinite population model cannot be demonstrated in any other way except via the process model, why bother with it at all? Strike five against the infinite population model.

p. 197, between “iid” and “cluster sampling”:

**Infinite Population Model**      A term that is sometimes used in place of the process model, simply to make people who like the term “population” feel more comfortable. The infinite population model can be justified only through the probabilistic process model, and the term “infinite population” is therefore unnecessary and should be avoided.

- p. 200, problem 7.11: ... a person's housing expense.  $\rightarrow$  ... that person's housing expense.
- p. 203, four lines below Table 8.1: ...that occur less frequently ...  $\rightarrow$  ...that will occur less frequently ...
- p. 203: The equation labels (8.1) and (8.2) should be aligned and given the same font size.
- p. 204: The equation labels (8.3) and (8.4) should be aligned and given the same font size.
- p. 207: Example 8.4 ends on the fourth line from the top. The sentence starting "The expected value..." is the start of the material that is outside the example.
- p. 208, line 7: "... gets closer to ..."  $\rightarrow$  "... tends to get closer to ..."
- p. 217, five lines below Table 8.8: ... because it is unknown...  $\rightarrow$  ... because  $p(y)$  is unknown...
- p. 239, line 13 from bottom: ...model is the stoplight example.  $\rightarrow$  ...model in the stoplight example.
- p. 251, Figure 251. A cosmetic improvement would be to remove the labels "Row1," "Row 2," etc.
- p.263, three lines below "Additivity Property ...": ... shown earlier ...  $\rightarrow$  ... shown above ...
- p. 266, sentence above third set-off derivation: "...that assumes both iid observations:"  $\rightarrow$  "...that requires independence as well as identically distributed observations:"
- p. 267, five lines into Section 10.3, remove the " $\wedge$ " from the  $\sigma_{xy}$ .
- p. 273, last two full paragraphs: The " $t$ " in the exponent of the expression for  $p(y)$  should be a " $y$ " (twice).
- p. 285, third line of Example 11.2: "than in it actually does."  $\rightarrow$  "than it actually does."
- p. 285, end of second-to-last paragraph in the Example 11.2 text: "... so the estimate is..."  $\rightarrow$  "... so the estimator is..."
- p. 286, First sentence of last paragraph: "...rolling the dice five time, computing ..."  $\rightarrow$  "...rolling the dice five times, then five more times, computing ..."
- p. 291, lines 8-9:  
 "... on average, the plug-in estimate tends to be too small."  $\rightarrow$   
 "... on average, the plug-in estimate is too small."
- p. 293, second line below first set-off equations:  $...=(1/n)\hat{\sigma}^2 \rightarrow ...=(1/n)\sigma^2$
- p. 305, 11.7F. ... estimator of  $\theta_1^2$ ?  $\rightarrow$  ... estimator of  $\theta^2$ ?
- p. 307, beginning of third paragraph: Some might prefer the phrasing "In addition to providing estimates

of parameters, the likelihood function also provides a set of values for your model's parameters that \*is\* consistent with your observed data"

p. 308, 13-14 from bottom:

"...you can see that the values of  $\theta$  are more consistent with the observed data,..." →

"...you can see the values of  $\theta$  that are more consistent with the observed data,..." →

p. 310, Figure 12.1 legend: ...  $\lambda$  (shown as "lambda" in the graphs). The ... → ...  $\lambda$  ("lambda"). The ...

p. 317, lines 8-9:

Looking at the vertical range of the contour plot in Figure 12.6, ... →

Looking at the horizontal range of the contour plot in Figure 12.7, ...

p. 317, lines 10-11: Looking at the horizontal range, ... → Looking at the vertical range, ...

p. 319, Definition of MLE: ... for all permissible  $\theta$ , then ... → ... for all permissible  $\theta \neq \hat{\theta}$ , then ...

p. 322, Example 12.6, third reason:

(By property of logarithms that  $\ln(e^x) = x$ ) →

(By property of logarithms that  $\ln(a^x) = x\ln(a)$  and  $\ln(e) = 1$ )

p. 333, Figure 12.23: Vertical axis label should be "Probability \*of\* success"

p. 335, Second reason in first derivation:

(By derivative properties of a constant is zero, property D1, D2, and D8) →

(By derivative properties D1, D2, and D8)

p. 338, 12.1C: ...find its maximum... → ...find the MLE...

p. 338, 12.1D:

Find the maximum of the log-likelihood function using an... →

Find the  $\hat{\pi}$  that maximizes the log-likelihood function using an...

p. 339, 12.5C: ...find its maximum ... → ...find the MLE ...

p. 339, 12.5D: Find the maximum of the ... → Find the MLE using the ...

p. 339, 12.8D: Report the Wald standard error ... → Report the Wald standard errors ...

p. 361, right below first displayed equation: "1s00" should be "100"

p. 363, ...five lines into Example 13.8:

“...make a reasonable proxy...” → “...are a reasonable proxy...”

p. 363, about two thirds of the way down the page:

...( $\mu$ ,  $\sigma$ ); given the  $n = 14$  observations, ... →

...( $\mu$ ,  $\sigma$ ) given the  $n = 14$  observations,

p. 370, problem 13.5B. Figure 12.5 → Figure 12.6 (twice)

p. 370, problem 13.8 problem statement:

You can turn this around to state that  $\theta | \hat{\theta} \sim N(\theta, \hat{\sigma}^2)$ , ... →

You can turn this around to state that  $\theta | \hat{\theta} \sim N(\hat{\theta}, \hat{\sigma}^2)$ , ...

p. 375, line 3:

In 95% of repeated samples of size  $n$  from the same process, ... →

In repeated samples of size  $n$  from the same process, ...

p. 375, second bullet point:

... (normal, Poisson, etc.). →

...(i.e., you often do not need to assume a normal, Poisson, or any other “named” distribution when you use frequentist methods.).

p. 384, line 2: Insert ) between each  $\sqrt{16}$  and the following comma (two insertions)

p. 385, lines 8-9 from bottom: “...these days you will see advanced statistical methods that often...” →  
“...these days you will often see advanced statistical methods that ...”

p. 387, Exercise 14.4: ... from Exercise 12.9. → ...from Exercise 12.9, and assume that both the  $x$  data and  $y$  data are produced as iid from their respective distributions.

p. 390, first sentence of third paragraph of Section 15.2:

You can equate results of interest to differences found in your data. →

You can equate “results of interest” to “differences found in your data.” →

p. 395, line 10: ...using the normal distribution, ... → ...using normal distributions, ...

p. 399, line 17: ... is a permanent fixture on the statistical landscape. → ... has been a pronounced fixture on the statistical landscape.

p. 400, last line: -0.0004514 → -0.0004495

p. 401, line4-5:  $0.00080521 - (-0.0004514) = 0.001257$  (or 0.1257%) →  $0.00080521 - (-0.0004495) = 0.001255$  (or 0.1255%)

p. 401, 402: Several similar slight changes as noted above.

p. 402, Figure 15.4, horizontal axis:

-0.0000 -0.0005 -0.0010 -0.0015 →

0.0000 0.0005 0.0010 0.0015

p. 407, Problem 15.2B: ... distribution in Exercise 15.2B... → ... distribution in Exercise 15.2A...

p. 408, Problem 15.4C: ... answer to Exercise 15.4C, ... → ... answer to Exercise 15.4B, ...

p. 410, line 2 from bottom: ... be not mysterious ... → .. not be mysterious ...

p. 417, First sentence of last paragraph:

... standard random normal random variables ... → ... standard normal random variables ...

p. 418, 3 lines above 1<sup>st</sup> displayed equation: ... your  $N(\mu, \sigma^2)$ . → ... your  $N(\mu, \sigma^2)$  model.

p. 422, second line: ... Equations 16.3 and 16.4... → ... Equations 16.4 and 16.5...

p. 425, first line below Equation 16.10: ... iid  $N(\mu, \sigma^2)$  is valid ... → ... iid  $N(\mu, \sigma^2)$  model is valid ...

p. 429, bottom of page, the last reason on the right hand side:

... that the  $\bar{Y}$ 's and  $\hat{\sigma}_{pooled}$  are independent ... →

... that  $\bar{Y}$  and  $\hat{\sigma}_{pooled}$  are independent ...

p. 432, line 5: ... given in Chapter 15 approaches is ... → ... given in Chapter 15 is ...

p. 435, line 2 of Section 16.9: ... functions iid standard ... → ... functions of iid standard ...

p. 438, first two lines:

“...the less likely it is that the differences between sample means of the groups can be explained by chance alone.” →

“...the more difficult it is to explain the differences between sample means as resulting from chance alone.”

p. 452, last line of third paragraph: The  $\mu_i$  is too high: It should be typeset in line with the rest of the sentence.

p. 456, the given reason in the second of the four derivations:

(By property of logarithms) → (By properties of logarithms)

p. 457, bottom line: There seems to be a printing glitch that cuts off the bottom of the first  $\bar{y}$ .

p. 459, top line:

Recall from Example 12.2 that the likelihood function for the exponential sample is ... →

From Example 12.2, the log likelihood function for the exponential sample is ...

p. 461, derivation, first and third lines: Delete the “{” bracket.

p. 464, bottom line:  $\hat{\beta}_4 \rightarrow \hat{\beta}_k$

p. 465, top line:  $\beta_4 \rightarrow \beta_k$

p. 470, denominator of last displayed equation:  $n - 1 \rightarrow n - g$

p. 473, first two displayed equations: The “=0” and “=1” in the subscripts of the summation symbols should be typeset slightly higher, aligned with y, not with the subscript of y.

p. 476, second displayed equation: the “ $f_i$ ” in the numerator should be “ $f_j$ ”.

p. 476, first set-off derivation, last line: Insert a “)” just before the “}” so it reads “...+ $f_6 \ln(1/6)$ ))”

p. 476, second and third lines below the first set-off derivation: The line break from “ $f_i$ ” at the end of the second line to “ $\ln(f_i/n)$ ” at the beginning of the third line is awkward. The expression “ $f_i \ln(f_i/n)$ ” should be kept contiguously intact.

p. 479, lines 2-3: “Starting”  $\rightarrow$  “Start”; “chi-squared statistic”  $\rightarrow$  “chi-squared statistic:”

p. 479, two lines above second-to-last displayed equation: “Supposing”  $\rightarrow$  “Suppose”

p. 479, one line above last displayed equation: “Thus”  $\rightarrow$  “Then”

p. 479, four lines from bottom: “So the last term in the summand...”  $\rightarrow$  “So the last summand...”

p. 480, last displayed equation:  $\hat{\pi}_{ji} = \frac{f_{ij}}{f_i} \rightarrow \hat{\pi}_{ji} = \frac{f_{ij}}{f_i}$

p. 481, before second-to-last displayed equation: “... conditional likelihoods as shown in the following:”  $\rightarrow$  “...conditional likelihoods, so the log likelihood is given by:”

p. 482, first displayed equation:  $f_j \rightarrow f_{.j}$

p. 482, the equation in Example 17.10, middle of page:

$$\chi^2 = 5 \times \ln(5/1.09) + 1 \times \ln(1/0.91) + 0 \times \ln(0/1.27) + \dots + 3 \times \ln(3/0.76) = 43.11$$

$\rightarrow$

$$\chi^2 = 2\{5 \times \ln(5/1.09) + 1 \times \ln(1/0.91) + 0 \times \ln(0/1.27) + \dots + 3 \times \ln(3/0.76)\} = 43.11$$

p. 482, 12 from bottom: “..., such permutation and ...”  $\rightarrow$  “..., such as permutation and ...”

p. 484, first sentence in Example 17.11: There is a bad line break at the end of the line. The expression  $\sum f_j \ln(f_j / n)$  should be kept intact.

p. 484, five lines above Figure 17.6: Remove the extra space between “-” and “148.60”.

p. 485, definition of “Covariate”: ... deterministically related to another variable  $Y$  ... →  
...deterministically related to the conditional mean (given  $X = x$ ) of another variable  $Y$  ...

p. 487, second description in the right column, line 5: The restricted ... → A restricted ...

p. 489, 17.13F:

Calculate the Pearson chi-squared test and ... →

Calculate the Pearson chi-squared statistic and ...

p. 494, lines 2-3: There is a missing hyphen in the line break. “exam ple” → “exam-ple”

p. 494, second equation in the derivation at the bottom of the page: There is a bad line break in the equation. The expression  $50/(500/96^{1/2})$  should be kept intact.

p. 498, lines 10-11 from bottom:

“...could be anything,  $N(311, 4.5^2)$ , or ...” →

“...could be  $N(311, 4.5^2)$ , or ...”

p. 507, third full paragraph. All numbers starting with “2” should start with “3” (7 changes).

p. 518, two lines below first displayed equation:

...pooled standard deviation calculated from ... →

...pooled variance estimate calculated from ...

p. 520, first flush left sentence: “If you are ... just sample ...” → “If you are ... just sampling ...”