



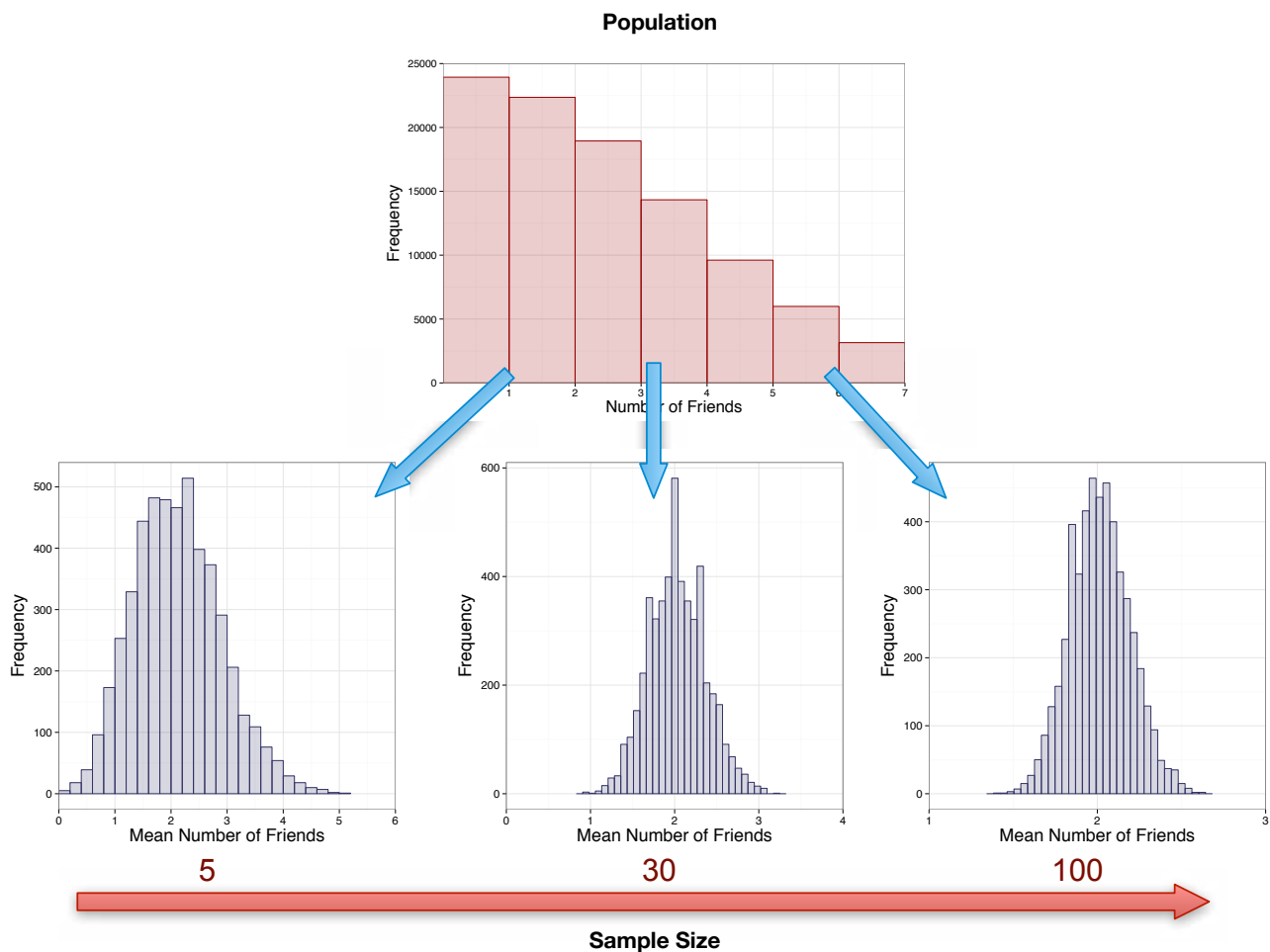
DSUS (4th Edition) Errata

Corrected After The First Print Run:

- P. 58: equation 2.11 for the upper boundary of the confidence interval should have a plus sign rather than a minus (bloody typesetters ...):

$$\text{upper boundary of confidence interval} = \bar{X} + (t_{n-1} \times SE)$$

- P. 162: The first task references data from Chapter 2; that should be Chapter 3 as per the reference to Table 3.1.
- P. 169-171 Figure 5.5, ought to look like this (x-axis scales on the lower part of the diagram were wrong):



- P. 118 Task 4, the table ought to look like this (the columns labelled *Distance* and *Time* under 'Female' are the wrong way round in the book):

Male		Female	
Distance	Time	Distance	Time
15	0.16	22	1.40
30	0.40	140	1.81
37	1.36	160	1.96
65	1.99	183	3.02
103	3.61	245	4.82

- P. 222: the K-S tests are reported with the wrong significance values (I mistakenly put the *p*-values for the Shapiro-Wilk test). The paragraph should read as follows (corrections in bold):



The tables in Output 6.1 confirm these observations: for the Sunday data the distributions for ecstasy, $D(10) = 0.28$, $p = .03$, appears to be non-normal whereas the alcohol data, $D(10) = 0.17$, $p = .20$, were normal; conversely, for the Wednesday data, although the data for ecstasy were normal, $D(10) = 0.24$, $p = .13$, the data for alcohol were significantly non-normal, $D(10) = 0.31$, $p = .009$.

- P. 265: Equation at the bottom of the page, second step: the fifth value in the numerator should be -1.4 instead of 0.4 (value in red below):

$$\begin{aligned} \text{covariance}(x, y) &= \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{N - 1} \\ &= \frac{(-0.4)(-3) + (-1.4)(-2) + (-1.4)(-1) + (0.6)(2) + (2.6)(4)}{N - 1} \\ &= \frac{1.2 + 2.8 + 1.4 + 1.2 + 10.4}{4} \\ &= \frac{17}{4} \\ &= 4.25 \end{aligned}$$

- P. 374: Output 9.3 is incorrect, and this has implications for the preceding text. This section should read as follows

The output from the independent t -test contains only three tables (two if you don't opt for bootstrapping). The first table (Output 9.3) provides summary statistics for the two experimental conditions (if you don't ask for bootstrapping this table will be a bit more straightforward). From this table, we can see that both groups had 12 participants (row labelled N). The group who had no cloak, on average, performed 3.75 mischievous acts with a standard deviation of 1.913. What's more, the standard error of that group is 0.552 ($SE = 1.913/\sqrt{12} = 1.913/3.464 = 0.552$). The bootstrap estimate of this SE is a little smaller at .54, and the bootstrapped confidence interval for the mean ranges from 2.67 to 4.75. For those that were given an invisibility cloak, they performed, on average, 5 acts, with a standard deviation of 1.651, a standard error of 0.477 ($SE = 1.651/\sqrt{12} = 1.651/3.464 = 0.477$). The bootstrap estimate of this SE is 0.47, and the confidence interval for the mean ranges from 4.14 to 5.86. Note that the confidence intervals for the two groups overlap, implying that they might be from the same population.

			Statistic	Bootstrap ^a			
				Bias	Std. Error	BCa 95% Confidence Interval	
Cloak of invisibility						Lower	Upper
Mischievous Acts	No Cloak	N	12				
		Mean	3.75	-.01	.54	2.67	4.75
		Std. Deviation	1.913	-.114	.352	1.320	2.226
		Std. Error Mean	.552				
Cloak	N	12					
	Mean	5.00	-.01	.47	4.14	5.86	
	Std. Deviation	1.651	-.089	.322	1.034	2.027	
	Std. Error Mean	.477					

a. Unless otherwise noted, bootstrap results are based on 1000 bootstrap samples

Output 9.3

- p. 629, last part of the page states '...of the 20 scores and...'. It actually refers to the 30 elements of DV1 in table 16.1.
- On p. 648 there is a reference to the group effects in Output 16.3. In the 7th line down in section 16.7.2. I refer to the degrees of freedom as 2 and 26. However, these are the df for the intercept effect, not the group effect. It should read 4 and 54 degrees of freedom.



Still Lurking in There:

- Page 32: For some reason the text doesn't match the Figure (1.14). I must have been having a funny half an hour. The text should read (changes in bold): "For example, if we look up our z-score of 2.6, we find that the smaller portion (i.e., the area above this value, or the blue area in Figure 1.14) is **.0047**, or put another way, only **0.47%**. I explained before that these areas relate to probabilities, so in this case we could say that there is only a **0.47%** chance that a suicide victim would be 70 years old or more. By looking at the larger portion (the area below 2.60) we get **.9953**, or put another way, there's a **99.53%** chance that a suicide victim was younger than 70 years old. Note that these two proportions add up to 1 (or 100%), so the total area under the curve is 1". [Paolo Vaccarino]
- Page 184 (5.3.2.2): the last but one paragraph, the first equation ' $-0.004/0.086 = 0.047$ ' should be ' $-0.004/0.086 = -0.047$ ' (note minus sign on right hand side). [Jiawei (Hugo) Sun]
- Chapter 10 Mediation (p 412-413). A paper in 2015 (Wen, Z., & Fan, X. (2015, February 9). Monotonicity of Effect Sizes: Questioning Kappa-Squared as Mediation Effect Size Measure. Psychological Methods. Advance online publication. <http://dx.doi.org/10.1037/met0000029>) has demonstrated that kappa-squared is not a good effect size measure - they recommend using P_M for simple mediation models. I would revise my advocacy of Kappa-squared in light of this paper. (Thanks to Kris Preacher for sending me this paper.)
- Page 830: "People learning statistics often wonder about "their" being a best way to do things..." 'Their' should be 'there'. [Kate Talsma]
- Glossary: Trimmed mean, the data example is wrong. It should read: **Trimmed mean**: a statistic used in many *robust tests*. It is a mean calculated using *trimmed data*. For example, a 20% trimmed mean is a mean calculated after the top and bottom 20% of ordered scores have been removed. Imagine we had 20 scores representing the annual income of students (in thousands), rounded to the nearest thousand: 0, 1, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 40 The mean income is 5 (£5000), which is biased by an outlier. A 10% trimmed mean will remove 10% of scores from the top and bottom of ordered scores before the mean is calculated. With 20 scores, removing 10% of scores involves removing the top and bottom 2 scores. This gives us: 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, the mean of which is 3.44. The mean depends on a symmetrical distribution to be accurate, but a trimmed mean produces accurate results even when the distribution is not symmetrical. There are more complex examples of robust methods such as the *bootstrap*.

Thanks to everyone spotting mistakes.