Instructor Solutions Manual

The Statistical Sleuth
A Course in Methods of Data Analysis

THIRD EDITION

Fred L. Ramsey
Oregon State University

Daniel W. Schafer
Oregon State University

BROOKS/COLE
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

BROOKS/COLE
CENGAGE Learning

ISBN-13: 978-1-133-59054-5
ISBN-10: 1-133-59054-3

Brooks/Cole
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at:
www.cengage.com/global

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Brooks/Cole, visit
www.cengage.com/brookscole

Purchase any of our products at your local college store or at our preferred online store
www.cengagebrain.com

Printed in the United States of America
1 2 3 4 5 6 7 16 15 14 13 12

Contents

Preface: To The Instructor

This solutions manual contains answers and sketches of solutions to all "computational exercises" and "data problems" that appear at the ends of the chapters in *The Statistical Sleuth.* The Student Solutions Manual contains identical information, but for selected "computational exercises" only. About half of the computational exercises answers are provided in the Student Solutions Manual.

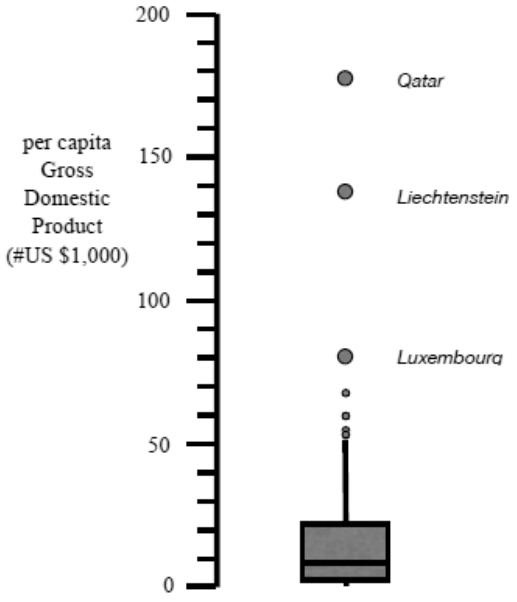
The "data problems" are, of course, important for practical experience at real data analysis and communication of statistical results. Many of these are quite hard—mainly because real data problems can be quite hard. We provide sketches of solutions to the data problems here, but wish to point out that there is often more than one correct approach. We hope that students use the "Statistical Conclusions" sections at the end of each case study in the book as templates for their own wording of results.

We will periodically provide updates and corrections on the web site www.statisticalsleuth.com. There are instructions there for joining our mailing list so that you may receive any updates or news that we believe worthy of broadcasting. You may contact us by e-mail at: ramsey@stat.orst.edu orschafer@stat.orst.edu.

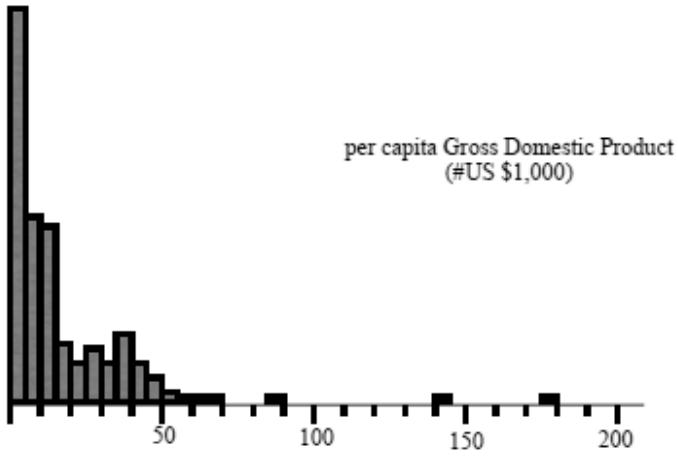
Chapter 1: Drawing Statistical Conclusions

1.16 Gross Deomestic Product (GDP) Per Capita.

a



c



1.17 The difference between averages (A − B) in the observed outcome is 78.00 − 62.67 = +15.33 points. In the list that follows, there are three outcomes (nos. 1, 34, and 35) that have a difference as large or larger in magnitude as the observed difference. The two-sided *p-value* is therefore 3/35 = 0.0857.

Outcome No.	Guide A	A Average	Guide B	B Average	(A − B) Difference
1	53, 64, 68, 71	64.00	77, 82, 85	81.33	−17.33
2	53, 64, 68, 77	65.50	71, 82, 85	79.33	−13.83
3	53, 64, 68, 82	66.75	71, 77, 85	77.67	−10.92
4	53, 64, 68, 85	67.50	71, 77, 82	76.67	−9.17
5	53, 64, 71, 77	66.25	68, 82, 85	78.33	−12.08
6	53, 64, 71, 82	67.50	68, 77, 85	76.67	−9.17
7	53, 64, 71, 85	68.25	68, 77, 82	75.67	−7.42
8	53, 64, 77, 82	69.00	68, 71, 85	74.67	−5.67
9	53, 64, 77, 85	69.75	68, 71, 82	73.67	−3.92
10	53, 64, 82, 85	71.00	68, 71, 77	72.00	−1.00
11	53, 68, 71, 77	67.25	64, 82, 85	77.00	−9.75
12	53, 68, 71, 82	68.50	64, 77, 85	75.33	−6.83
13	53, 68, 71, 85	69.25	64, 77, 82	74.33	−5.08
14	53, 68, 77, 82	70.00	64, 71, 85	73.33	−3.33
15	53, 68, 77, 85	70.75	64, 71, 82	72.33	−1.58
16	53, 68, 82, 85	72.00	64, 71, 77	70.67	+1.33
17	53, 71, 77, 82	70.75	64, 68, 85	72.33	−1.58
18	53, 71, 77, 85	71.50	64, 68, 82	71.33	+0.17
19	53, 71, 82, 85	72.75	64, 68, 77	69.67	+3.08
20	53, 77, 82, 85	74.25	64, 68, 71	67.67	+6.58
21	64, 68, 71, 77	70.00	53, 82, 85	73.33	−3.33
22	64, 68, 71, 82	71.25	53, 77, 85	71.67	−0.42
23	64, 68, 71, 85	72.00	53, 77, 82	70.67	+1.33
24	64, 68, 77, 82	72.75	53, 71, 85	69.67	+3.08
25	64, 68, 77, 85	73.50	53, 71, 82	68.67	+4.83
26	64, 68, 82, 85	74.75	53, 71, 77	67.00	+7.75
27	64, 71, 77, 82	73.50	53, 68, 85	68.67	+4.83
28	64, 71, 77, 85	74.25	53, 68, 82	67.67	+6.58
29	64, 71, 82, 85	75.50	53, 68, 77	66.00	+9.50
30	64, 77, 82, 85	77.00	53, 68, 71	64.00	+13.00
31	68, 71, 77, 82	74.50	53, 64, 85	67.33	+7.17
32	68, 71, 77, 85	75.25	53, 64, 82	66.33	+8.92
33	68, 71, 82, 85	76.50	53, 64, 77	64.67	+11.83
34	68, 77, 82, 85	78.00	53, 64, 71	62.67	+15.33
35	71, 77, 82, 85	78.75	53, 64, 68	61.67	+17.08

1.18 Outcomes will vary with different randomizations. See text Display 1.7

1.19 Coin flips will not divide the subjects in such a way that there is an exact age balance. However, it is impossible to tell prior to the flips which group will have a higher average age.

1.20 The randomization scheme suggested in problem 18 works. So would dealing five red and five black cards after shuffling. Once again it will not guarantee an exact age balance, but the group that gets the higher average is not predictable in advance of the randomization.

1.21 There is no computation involved. This is, however, a sobering exercise.

1.22

Intrinsic ($n = 24$)		*Extrinsic* ($n = 23$)
	5 | 04	
	6 | 1	
	7 |	
	8 |	
	9 |	
	10 | 9	
	11 | 8	
900	12 | 03	
6	13 |	
	14 | 8	
	15 | 0	
6	16 | 8	
52	17 | 2245	
2	18 | 577	
831	19 | 25	
653	20 | 7	
63	21 | 2	
621	22 | 1	
1	23 |	
30	24 | 0	
	25 |	
7	26 |	
	27 |	
	28 |	
7	29 |	
	30 |	

Legend: | 29 | 7 represents a creativity score of 29.7

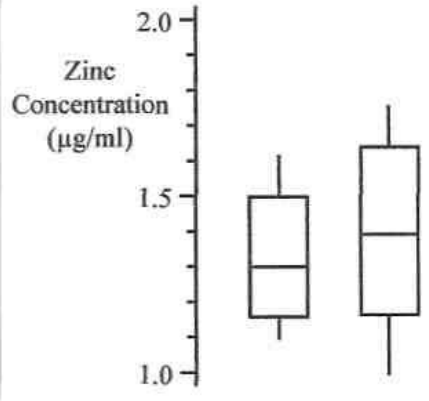
1.23 The box plot should look a bit like the stem and leaf diagram in exercise #22.

1.24 (Int,Ext): Medians are (20.4,17.2); Lower quartiles are (17.35, 12.0); Upper quartiles are (22.4, 19.2); IQRs are (5.05,7.2). There are no extreme points in either group.

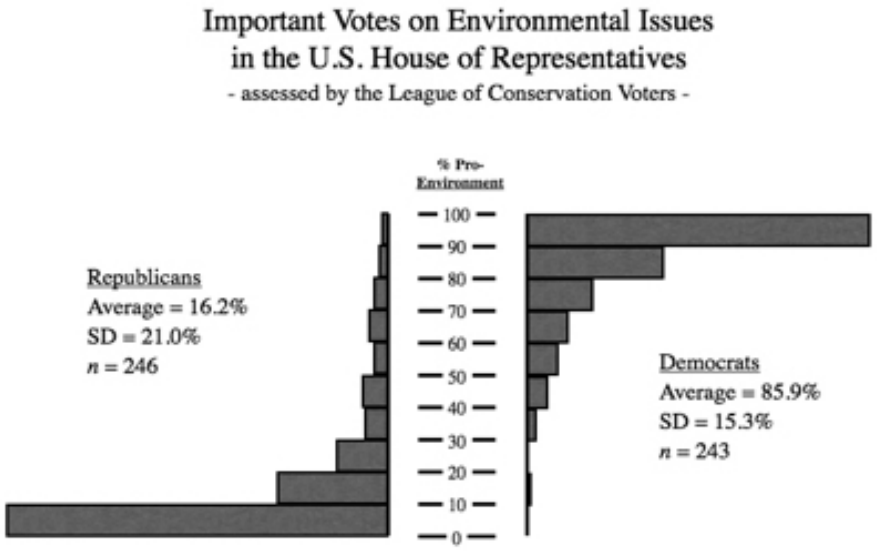
1.25

Group A ($n = 20$)			Group B ($n = 19$)
	10	03	
764210	11	3559	
9320	12	0	
10	13	349	
952	14	77	
9320	15	59	
2	16	8	
	17	1446	

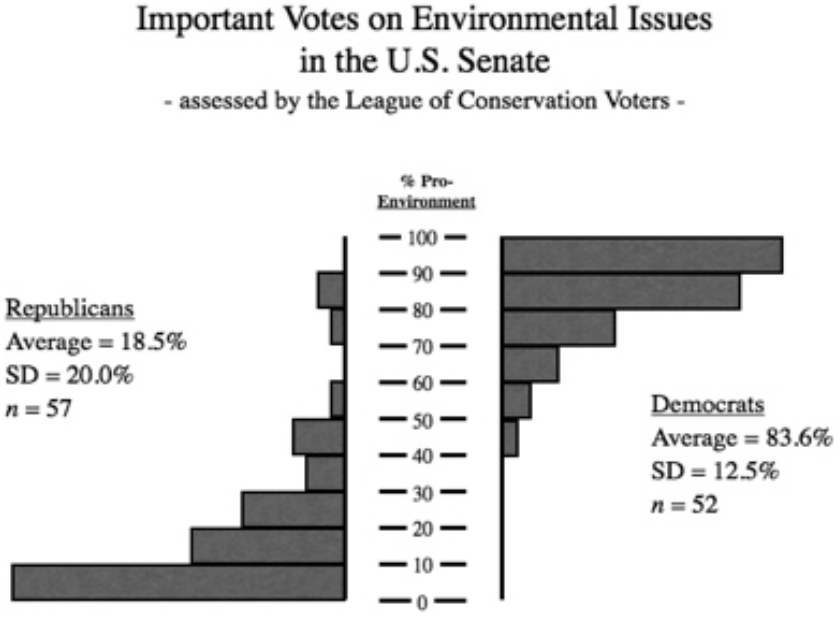
Legend: | 17 | 6 represents 1.76 µg/ml



1.26 Any picture tells the story. There is no need for a statistical test.



1.27 Again, any picture tells the story. There is no need for a statistical test.



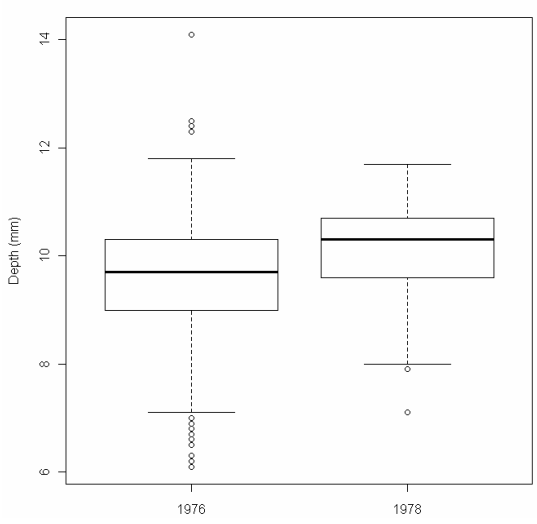
Chapter 2: Inference Using *t*-Distributions

2.12 **a** From 18 8 to 372 grams.
b From 203 to 357 grams. In both these, round *out,* which can be accomplished by rounding the halfwidth up before adding and subtracting.
c *t*-statistic = 6.00. The two-sided *p*-value is <.0001. (It is minuscule.)

2.13 **a** (Fish, Regular): Averages are (6.571, −1.143); SDs are (5.855, 3.185)
b Pooled SD = 4.713
c SE for difference = 2.519
d d.f. = 12; $t_{12}(.975) = 2.179$
e 95% CI from 2.225 to 13.203 mm
f *t*-stat = 3.062
g One-sided *p*-value = .005. Using the table in Appendix 2, locate the d.f. =12 line, and move across the line until the position of the *t*-statistic, 3.062. It is slightly larger than 3.055 so the table tells you that the one-sided *p*-value is slightly smaller than .005.

2.15 *t*-statistic = 9.32, with 174 d.f. Very convincing, indeed.

2.18 **The Grants' Complete Finch Beak Data.**
a



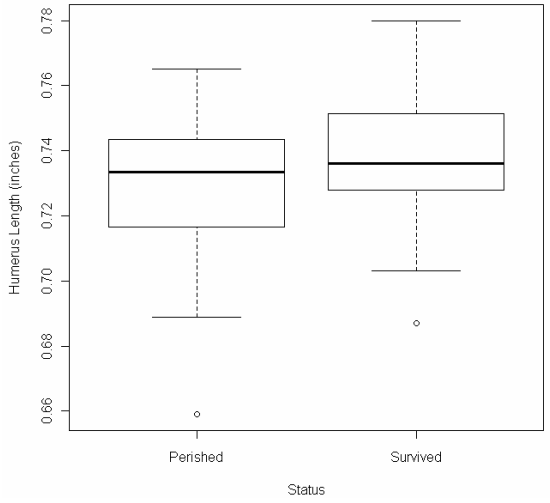
b one-sided *p*-value = 1.617e−06 (equal variance version of two-sample *t*-test)
c two-sided *p*-value = 3.233e−06 (equal variance version of two-sample *t*-test)
d Estimate of 1978 mean minus 1976 mean: 0.54mm; 95% confidence interval: 0.31 to 0.76mm
e Some of the finches may be in both samples or some of the finches in the 1978 sample may be offspring of some in the 1976 sample.

2.19 **a** Average = −1.14; SD = 3.18; d.f. = 6.
b SE = 1.20
c 95% CI: from −4.09 to 1.80
d *t*-statistic = −0.95; two-sided *p*-value = .38. [Using the table: 0.906 < 0.95 < 1.134, so 0.95 is between the 80th and the 85th percentiles. The one-sided *p*-value is therefore between 1−0.80 = 0.20 and 1−0.85 = 0.15, and the two-sided *p*-value is between 0.40 and 0.30 (by doubling).]

2.20 *t*-statistic = 2.97; two-sided *p-value* = 0.025. [Using the table: 2.612 < 2.97 < 3.143, so the one-sided *p*-value is between 1−.99 = .01 and 1−.98 = .02. The two-sided *p*-value is between .02 and .04 (by doubling).] The typical reduction is about 6.6 mm of mercury. The 95% confidence interval is from 1.15 to 11.99 mm.

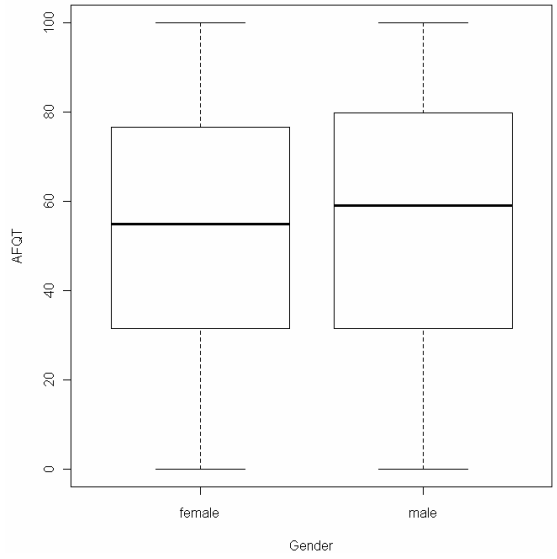
DATA PROBLEMS

2.21 **Bumpus Natural Selection Data**.

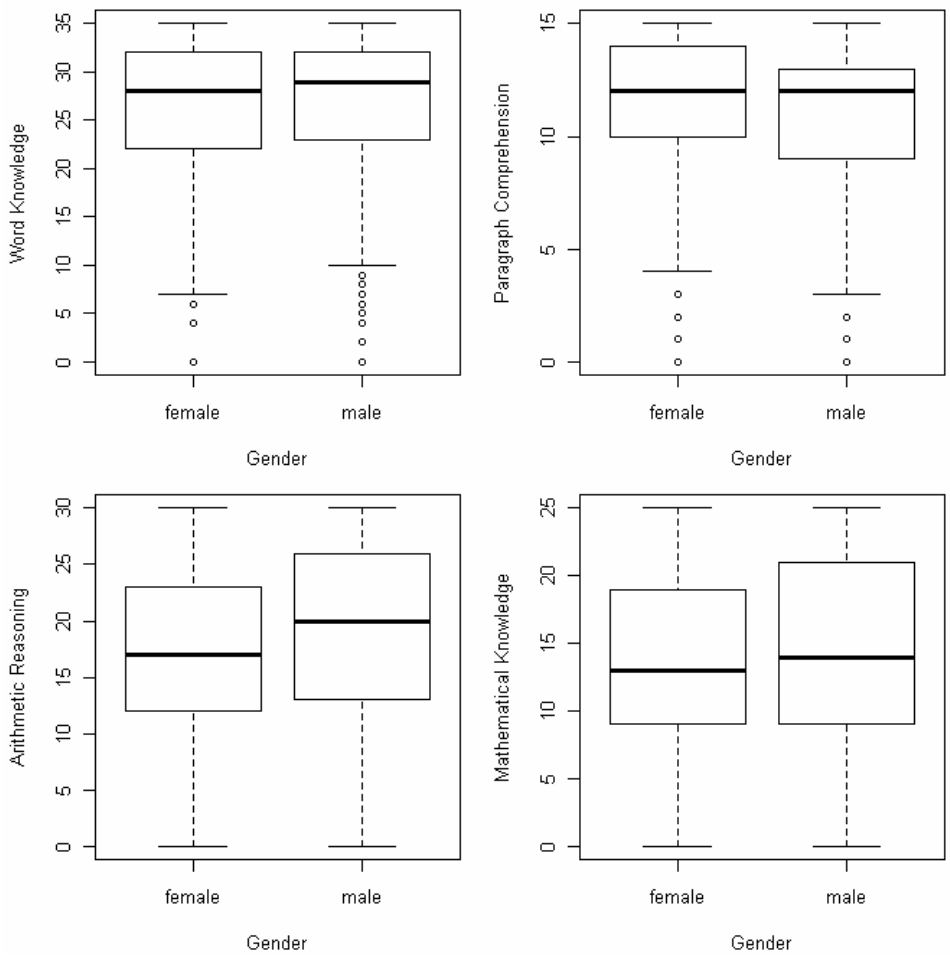


These data provide suggestive but inconclusive evidence that the distribution of humerus lengths differed in the populations of sparrows that perished and survied (2-sided *p*-value = 0.08 from a two-sample *t*-test). The mean for the population that survived is estimated to exceed the mean for the population that perished by 0.0101 inches (95% confidence interval: −0.0214 to 0.0013 inches).

2.22 **Male and Female Intelligence**.



These data provide suggestive but inconclusive evidence that the distribution of AFQT scores for males and females differ (2-sided *p*-value = 0.062). The male mean is estimated to exceed the female mean by 2.04 percentage points (95% confidence interval for male excess: − 0.10 points to 4.18 percentage points).



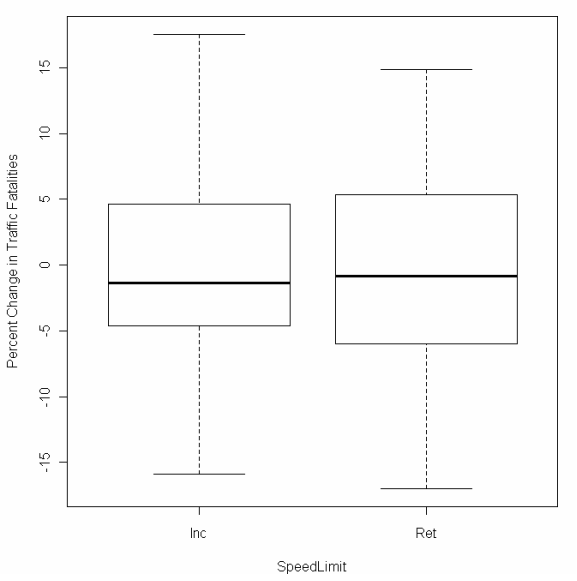
The data provide no evidence of a gender difference on Word Knowledge test scores (2-sided p-value = 0.94). The female mean is estimated to exceed the male mean by 0.02 percentage points (95% confidence interval −0.57 to 0.52 percentage points).

The data provide convincing evidence of a gender difference on Paragraph Comprehension test scores (2-sided p-value = 0.0000045). The female mean is estimated to exceed the male mean by 0.57 percentage points (95% confidence interval 0.32 to 0.81 percentage points).

The data provide convincing evidence of a gender difference on Arithmetic Reasoning test scores (2-sided p-value = 0.0000000000004). The male mean is estimated to exceed the female mean by 2.04 percentage points (95% confidence interval 1.49 to 2.58 percentage points).

The data provide strong evidence of a gender difference on Mathematical Knowledge test scores (2-sided p-value = 0.002). The male mean is estimated to exceed the female mean by 0.75 percentage points (95% confidence interval 0.27 to 1.24 percentage points).

2.23 Speed Limits and Traffic Fatalities.



The data provide no evidence that the mean percentage increase was higher in states that increased their speed limit than in states that didn't (1-sided p-value = 0.44). The mean percentage increase in traffic fatalities in states that increased their speed limit was estimated to exceed the mean in states that didn't by 0.39 percentage points (95% confidence interval: −4.4 to 5.2 percentage points).

Chapter 3: A Closer Look at Assumptions

3.20 **a** average In-State tuition = $11,600
log(average In-State tuition) = 9.3588
average log(In-State tuition) = 8.6606

b median In-State tuition = $5,000
log(median In-State tuition) = 8.5172
median log(In-State tuition) = 8.5172

c median Out-of-State tuition = $30,000
(median Out)/(median In) = 6.0
median(Out/In) = 3.0
median[log(Out) − log(In)] = 1.0986 = log(3)

3.21 **a** One-sample t-test on differences (observed − expected) for the subset of umpires whose lifetimes were not censored (Censored = 0): t-stat = −0.987, df = 194, p-value = 0.32 (1-sided p-value = .16). A 95 percent confidence interval for mean life length minus expected life length: −1.6 years to 0.54 years.

b This might be a problem if the ones for whom data were unavailable tended to have died young. In any case, the available sample is not a random sample from the population of all umpires.

c This is a considerable problem since with the given sampling routine we are more likely to sample umpires who died young than umpires who died old. For this reason the t-test based on the uncensored lifetimes is not a good idea here. (It is also inappropriate to insert artificial death times for the censored group; more sophisticated techniques of *survival analysis* would be needed.)

3.22 **a** at 26 kV: 1.756 7.365 7.751
at 28 kV: 4.231 4.685 4.703 6.055 6.973

b At (26 kV, 28 kV): Averages = (5.624, 5.329), SDs = (3.355, 1.145), *n*'s = (3, 5). Difference in averages = 0.295.

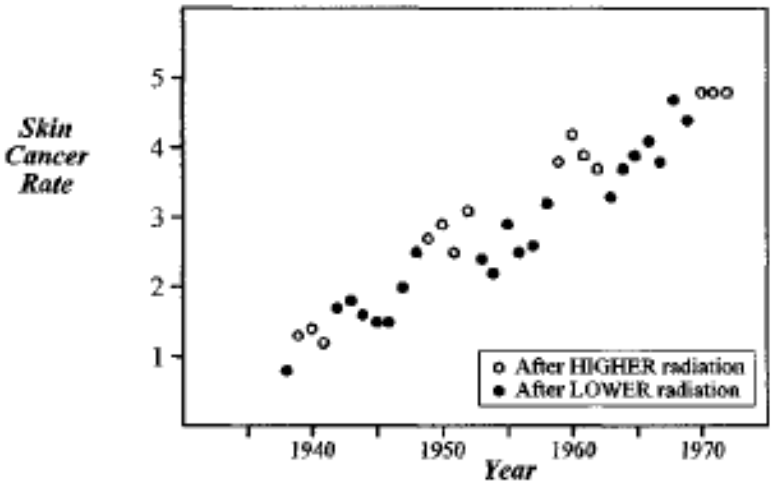
c exp(0.295) = 1.343 estimates the multiplicative effect on time to breakdown of changing voltage level from 28 kV to 26 kV.

d 95% CI goes from −4.138 to 3.549. Antilogs: from 0.016 to 34.8. The multiplicative effect of raising the voltage from 26 kV to 28 kV is estimated to be between 0.016 and 34.8 (95% confidence interval). So, you expected better precision?

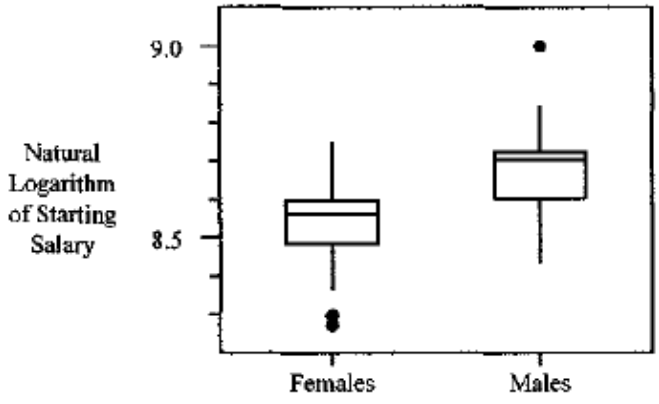
3.23 Refer to Display 3.10.

a Yes. One should expect the rates to follow a time series where serial correlation is present.

b Following is a picture that puts them both together. There is a problem: there is a steady increase, or 'trend', in the series. There is also a somewhat cyclic behavior. The trend and (possibly) the cyclic behavior are most likely unrelated to solar radiation, but they will have a strong influence on the comparison because more of the 'after higher' values fall in the later years.



3.24 **a** The box plots look like this:



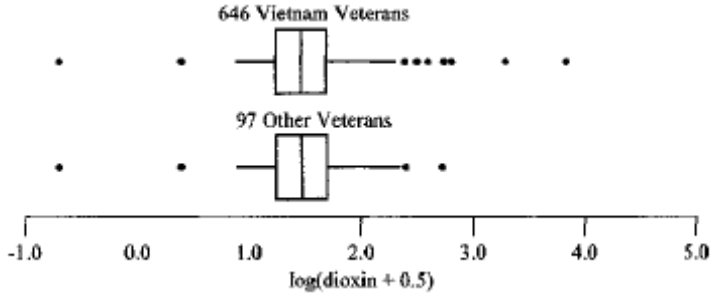
b Based on a t-statistic of 6.17 with 91 degrees of freedom, the one-sided p-value is <.0001.

c The 95% confidence interval for the ratio of median starting salaries (M/F) is from 1.10 to 1.21. (A 95% confidence interval for the ratio of median starting salaries (F/M) is from 1/1.21 to 1/1.10, or .82 to .91.)

3.25 Use the computer. Refer to Display 3.6.

3.26 **Agent Orange.** Use log(dioxin + 0.5) as the response.

a

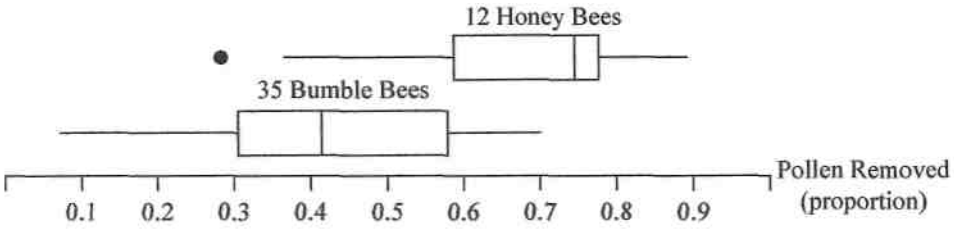


b Two-sided p-value = .3816

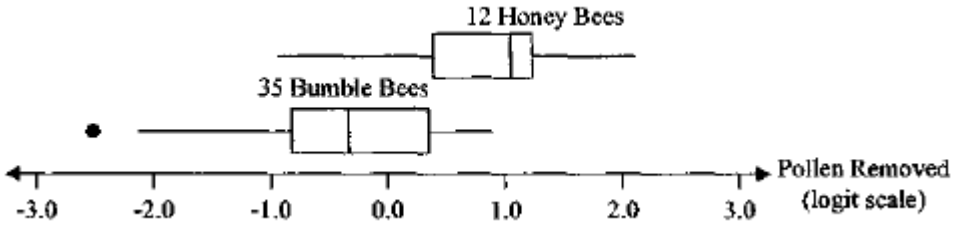
c (−0.0539, +0.1410) on the transformed scale converts to (0.9475, 1.1514) by anti-logs. These are bounds on the ratio of (median + 0.5) values (Vietnam to Other).

3.27

a (i)

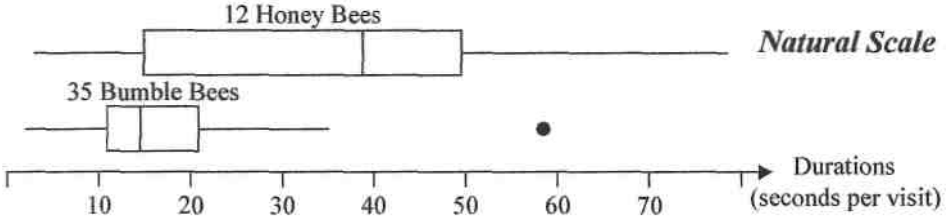


(ii)

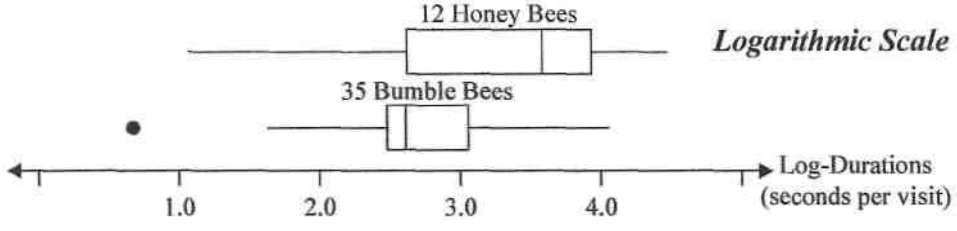


(iii) t = 3.85, two-sided p-value = .0004.

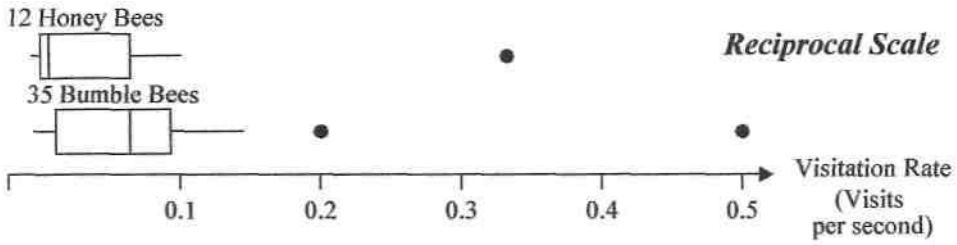
b (i)



(ii)



(iii)

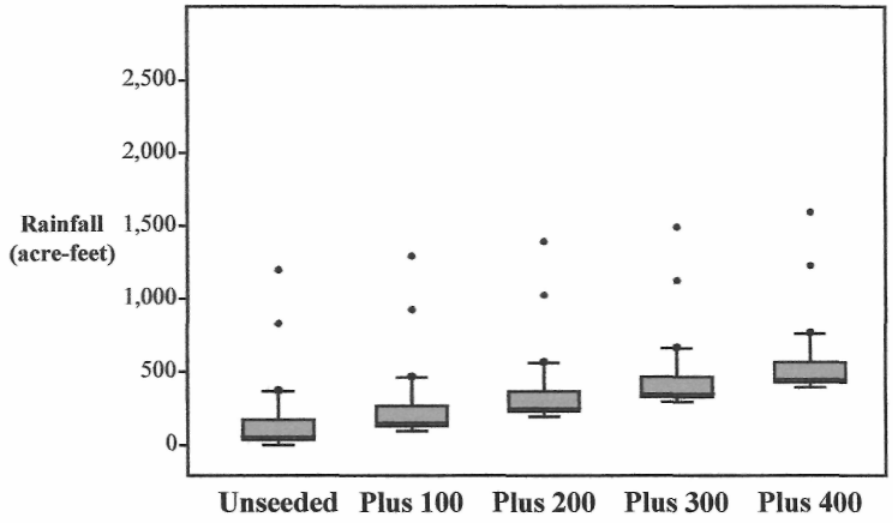


(iv) The log scale looks best (if not ideal). (v) Natural scale: from 10.0 to 29.4. Log scale: from 0.18 to 1.12. Reciprocal scale: from −0.081 to 0.031. (vi) All three are relatively easy to deal with. The reciprocal scale is a natural scale, but converting differences back would not give anything meaningful. (vii) It's very difficult.

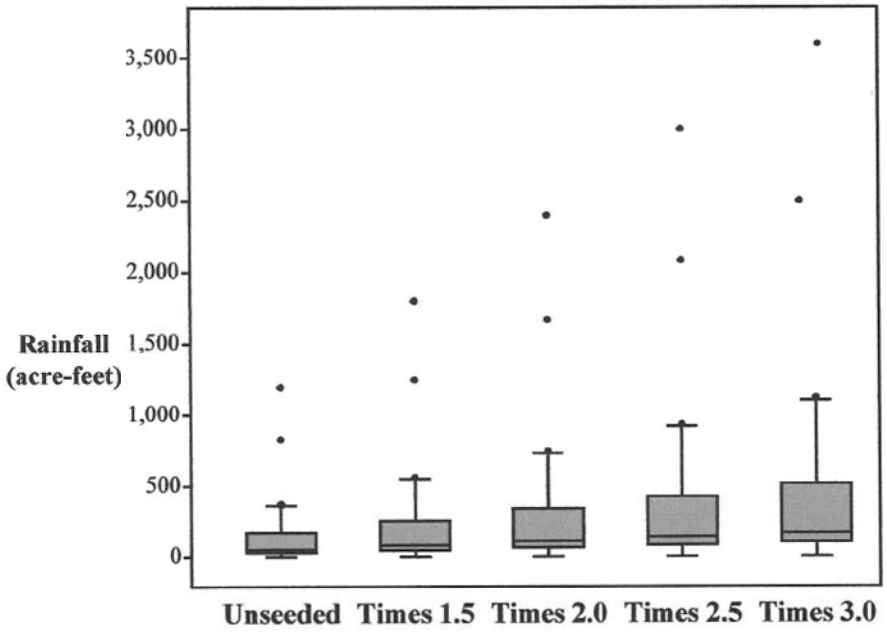
3.28 With all the data, the one-sided p-value is 0.0405; without the .659 value, the one-sided p-value is .0900. This is a fair swing; the evidence goes from suggestive to none.

3.29 **Cloud Seeding.**

a Additive changes to be unseeded distribution:



b Multiplicative (1.5 − 3.0, not 2 – 5) changes to the unseeded distribution:



c The data (Display 3.2) look more like the multiplicative effect of part **b**.

DATA PROBLEMS

3.30 **Education and Future Income.** The incomes span several orders of magnitude, so they are log-transformed. A test of the difference in average logs results in a two-sided p-value < 0.0001, providing convincing evidence of a real difference (subject to caveats concerning sampling). Upon back-transformation, it is estimated that students completing 16 years of education have 1.77 times (77% higher) the salary of students completing only 12 years. A 95% confidence interval on this multiplicative factor is from 1.60 (60% higher) to 1.96 (96% higher).

3.31 **Education and Future Income II.** Incomes are again log-transformed. A test of the difference in average logs results in a two-sided p-value = 0.165, providing no evidence of a real difference (subject to caveats concerning sampling). Upon back-transformation, it is estimated that students completing >16 years of education have 1.11 times (11% higher) the incomes of students completing only 16 years. A 95% confidence interval on this multiplicative factor is from 0.96 (4% lower) to 1.28 (28% higher).

3.32 **College Tuition.** Box plots or histograms of tuitions suggest log-transformation.

a This is a one-sample problem, with the response of log(Out/In). One-sample t-tools provide convincing evidence (two-sided p-value < 0.0001) that out-of-state tuitions exceed in-state tuitions in public schools. Back-transformation estimates that out-of-state tuition is 2.29 times (129% higher) than in-state tuition. The 95% confidence interval is from 2.01 to 2.61.

b This is a two-sample problem, again with log-transformed tuitions. The two-sided p-value ≤ 0.0001 provides convincing evidence of a difference between in-state tuitions of private and public schools. Back-transformation estimates that in-state tuition at private schools is 3.69 times what it is in public schools (95% confidence interval from 2.90 to 4.69).

c This is also a two-sample problem, and again on the transformed scale. The two-sided p-value = 0.0002 provides convincing evidence of a difference between out-of-state tuitions of private and public schools. Back-transformation estimates that out-of-state tuition at private schools is 1.61 times what it is in public schools (95% confidence interval from 1.27 to 2.05).

3.33 **Brain Size and Litter Size.** Pictures make the log scale an obvious choice. The evidence that brain sizes relative to body sizes are unequal is strong, but not convincing (two-sided p-value = .0512, from two-sample $t = 1.975$ with 94 d.f.). If relative brain weight (RBW) is l,000 × (brain weight/body weight), it is estimated that the median RBW among species with larger litter sizes (Š2) is 48.7% larger than the median RBW among species with smaller litter sizes (<2). A 95% confidence interval on this factor is (0.2% smaller, 121.7% larger).

Chapter 4: Alternatives to the *t*-Tools

4.14 **O-Ring study**. The one-sided *p-value* from the *t*-test is .0004, compared to .00989 from the permutation test.

4.15 One-sided *p*-value = 2/10 = .20.

4.16

Group 1	Group 2	Aver. Diff.
4,5,6	7,12	−4.50
4,5,7	6,12	−3.67
4,5,12	6,7	+0.50
4,6,7	5,12	−2.83
4,6,12	5,7	+1.33
4,7,12	5,6	+2.17
5,6,7	4,12	−2.00
5,6,12	4,7	+2.17
5,7,12	**4,6**	**+3.00**
6,7,12	4,5	+3.83

One-sided *p*-value = 2/10 = .20 ... again.

4.17 **O-Ring study.** (136 + 170 + 10 + 85 + 10 + 10) / 10,626 = 421/10,626 = .0396.

4.18

Treatment	Control	Rank Sum
1,2,3	4,5,6	6
1,2,4	3,5,6	7
1,2,5	3,4,6	8
1,2,6	3,4,5	9
1,3,4	2,5,6	8
1,3,5	2,4,6	9
1,3,6	2,4,5	10
1,4,5	2,3,6	10
1,4,6	2,3,5	11
1,5,6	2,3,4	12
2,3,4	1,5,6	9
2,3,5	1,4,6	10
2,3,6	1,4,5	11
2,4,5	1,3,6	11
2,4,6	1,3,5	12
2,5,6	1,3,4	13
3,4,5	1,2,6	12
3,4,6	1,2,5	13
3,5,6	**1,2,4**	**14**
4,5,6	1,2,3	15

One-sided *p-value* = 2/20 = 0.10.

4.19 **a** 0.1718
b Normal approximation
c Continuity correction
d *t*-test gives $p = .081$; *t*-test with removal gives $p = .180$; rank sum gave $p = .1718$.
e The rank sum test is valid AND it uses all the data.

4.20 **a** **Trauma and metabolic expenditure.**
b Sum = $T = 82$
c $n_1 = 7$, $n_2 = 8$, $\mu_R = 8.00$, $\sigma_R = 4.4681$, Mean(T) = 56.00, SD(T) = 8.6333, $Z = 3.012$
d One-sided p-value = .0013.

4.21 **Trauma and metabolic expenditure.** $Z = 2.95369$; two-sided p-value = .0314

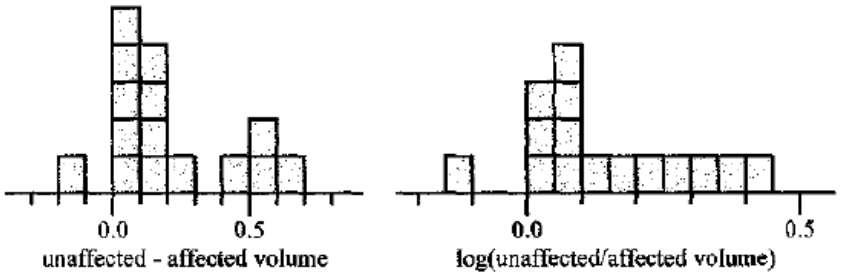
4.22 **Trauma and metabolic expenditure. (1.9, 16.8)**

4.23 **Motivation and creativity.** Two-sided p-value = .00643, compared to .00537.

4.24 **Motivation and creativity,** (1.00, 6.60), compared to (1.29, 7.00). (The former is based on the randomization test.)

4.25 **Guinea pig lifetimes.** CI: (39.59, 165.81), based on Welch's t with 97 d.f. The halfwidth is 63.11 and the critical t-multiplier is 1.9847. $SE_W = 31.80$ makes $t_w = 3.23$, giving a two-sided $p = .0016$. No. It looks like something else is involved.

4.26 **Schizophrenia Study, a** and **b.** There is mild skewness on the natural scale. The log scale is (marginally) better.



c On the log-scale, $t = 3.20$, for a two-sided $p = .0065$. On the 'natural' scale, the p-value is .0061, so there is virtually no difference.

d Estimate = 0.1285; 95% CI:(0.0423, 0.2147), which translates back to the original scale as ... estimate = 1.14; 95% CI: (1.04, 1.24).

4.27 **Schizophrenia Study.** Two-sided p-value = .00452, from signed rank test on the log(ratio) values. On the straight difference scale, the signed rank gives .00208 ... close. It is not particularly apparent.