Answer Key

1. **The following data are used on problems 1 and 2.**

   According to the City of Oxford, the speed limit on West Jackson Avenue is 30 m.p.h., while the mean speed (μ_x) of all traffic on the road is 36 m.p.h., with a standard deviation (σ_x) of 15 m.p.h.

   Approximately what percentage of traffic on the road obeys the speed limit? (i.e., what percentage of the traffic travels less than 30 m.p.h.)

   Answer:

   $$Z = \frac{X - \mu_x}{\sigma_x} = \frac{30 - 36}{15} = \frac{-6}{15} = -0.40$$

   Since we are interested in the area left of this value of Z, we obtain the small part value for $Z = 0.40$ in Appendix 1. We get a small part of .3446. So, 34.46% of traffic travels less than 30 m.p.h.

2. Approximately what is the 80th percentile speed of traffic on the road? (i.e. What speed does 80% of the traffic drive slower than?)

   Answer: We are interested in the big part that corresponds as close as possible to .8. In Appendix 1, the Z that corresponds to .8000 is approximately $Z = 0.84$.

   $$X = Z\sigma_x + \mu_x = (0.84)(15) + 36 = 48.6 \text{ m.p.h.}$$

   If you used 0.85 for Z, that was also considered correct; in that case, the answer obtained would have been 48.75 m.p.h.

3. What is the **probability** of correctly selecting the top three (3) finishers in order at the Brickyard 400, if 43 cars are in the race and all of the cars have an equal chance of winning? Hint: order matters.

   Answer:

$$nPm = \frac{N!}{(N-M)!} = \frac{43!}{(43-3)!} = \frac{43!}{40!} = 43 \times 42 \times 41 = 74046.$$

   Since there are 74046 possible permutations, and we are interested in just one possibility, $p = 1/74046 = .00001351$.

4. What is the probability of getting two heads out of eight flips of a balanced coin? Hint: order doesn't matter.

   Answer:

$$\begin{aligned} nCm &= \frac{N!}{M!(N-M)!} = \frac{8!}{2!(8-6)!} = \frac{8!}{2!6!} = \frac{8 \times 7}{2 \times 1} = 4 \times 7 = 28. \\ k^n &= 2^8 = 256 \\ p &= \frac{28}{256} = .109375 \end{aligned}$$

5. What is the probability of getting **four** or fewer heads out of eight flips of a balanced coin?

   Answer:

$$\begin{aligned} \text{Cumulative outcomes} &= \frac{8!}{4!(8-4)!} + \frac{8!}{3!(8-3)!} + \frac{8!}{2!(8-2)!} + \frac{8!}{1!(8-1)!} + \frac{8!}{0!(8-0)!} \\ &= 70 + 56 + 28 + 8 + 1 = 163. \\ p &= \frac{163}{256} = .63671875 \end{aligned}$$

6. The mean weight (μ_x) of members of the Ole Miss baseball team is 180 pounds, with a standard deviation (σ_x) of 30 pounds. We selected 14 players and their mean weight ($\bar{x}$) was 205 pounds. Construct a 95% confidence interval around the sample mean weight.

   Answer: Since $\alpha = .05$, $Z_{\alpha/2} = 1.96$.

$$\mu_x = \bar{x} \pm Z_{\alpha/2}\frac{\sigma_x}{\sqrt{n}} = 205 \pm (1.96)\frac{30}{\sqrt{14}} = 205 \pm (1.96)\frac{30}{3.74} = 205 \pm (1.96)(8.02) = 205 \pm 15.71$$

$$\rightarrow 189.29 < \mu_x < 220.71$$

7. Given the values provided in the problem above, test whether or not the sample is representative of the team; i.e., test whether $\mu_x = \bar{x}$, with a 99% confidence level.

   Answer: Since $\alpha = .01$, $Z_{\text{crit}} = 2.58$.

$$Z_{\text{ob}} = \frac{\bar{x} - \mu_x}{\sigma_x / \sqrt{n}} = \frac{205 - 180}{30/\sqrt{14}} = \frac{25}{30/3.74} = \frac{25}{8.02} = 3.118.$$

   Since $Z_{\text{crit}} < Z_{\text{ob}}$, we reject the null hypothesis and conclude that the sample mean and population means differ.

8. We surveyed 16 undergraduate students and found that, on average, their cars consumed 4 gallons of gasoline per week ($\bar{x}$), with a standard deviation (s_x) of 1.2 gallons. Construct a 99% confidence interval for μ_x around this sample mean.

   Answer: $\alpha = .01$ and df $= 16 - 1 = 15$, so $t_{\alpha/2} = 2.947$, from Appendix 2.

$$\mu_x = \bar{x} \pm t_{\alpha/2}\frac{s_x}{\sqrt{n}} = 4 \pm (2.947)\frac{1}{\sqrt{16}} = 4 \pm (2.947)\frac{1.2}{4} == 4 \pm (2.947)(0.3) = 4 \pm 0.8841$$
$$\rightarrow 3.1159 < \mu_x < 4.8841$$

9. Using the data provided in the previous problem, test whether or not the experience of our sample of students is typical of the population at large, given that the average car (μ_x) consumes 4.275 gallons of gasoline per week (i.e. test whether $\mu_x = \bar{x}$). Use a 95% confidence level.

   Answer: Since $\alpha = .05$ and df $= 16 - 1 = 15$, $t_{\text{crit}} = 2.131$.

$$t_{\text{ob}} = \frac{\bar{x} - \mu_x}{s_x / \sqrt{n}} = \frac{4 - 4.275}{1.2/\sqrt{16}} = \frac{-0.275}{0.3} = -0.916\overline{6}$$

   Since $t_{\text{crit}} > |t_{\text{ob}}|$, we fail to reject the null hypothesis and conclude that the sample and population means are not significantly different.

10. In November 2000, another enterprising undergraduate conducted a survey of attitudes toward the presidential candidates in the 2000 presidential election on the University of Mississippi campus. She found the following results:

Group	Greeks (1)	Non-Greeks (2)
n	46	76
Mean evaluation ($\bar{x}$) of Ralph Nader	36.2	28.7
Standard deviation (s_x)	4.2	10.3

Test the null hypothesis that $\bar{x}_1 = \bar{x}_2$; that is, test whether there was a significant difference between greek and non-greek evaluations of Ralph Nader. Use a 95% confidence level.

Answer: Since $\alpha = .05$ and df $= 46 + 76 - 2 = 120$, $t_{\text{crit}} = 1.9800$ (from Appendix 2).

$$
\begin{aligned}
s_{\text{pooled}} &= \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2} = \frac{45(4.2)^2 + 75(10.3)^2}{120} = \frac{45(17.64) + 75(106.09)}{120} \\
&= \frac{793.8 + 7956.75}{120} = 8750.55/120 = 72.92 \\
t_{\text{ob}} &= \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_{\text{pooled}}(1/n_1 + 1/n_2)}} = \frac{36.2 - 28.7}{\sqrt{72.92(1/46 + 1/76)}} = \frac{7.5}{\sqrt{72.92(0.0349)}} \\
&= \frac{7.5}{\sqrt{2.545}} = 7.5/1.595 = 4.70
\end{aligned}
$$

Since $t_{\text{crit}} < t_{\text{ob}}$, we conclude that there is a significant difference between greek and non-greek evaluations of Ralph Nader.

The number of vulgar T-shirts sold on Bourbon Street (New Orleans) during Mardi Gras by six vendors was tracked in 2001 and 2002. The data is presented below:

11.

Vendor	Sold	
Name	2001	2002
Bob	26	32
Steve	37	32
Maria	44	48
Louise	37	47
Jack	17	29
Henry	43	17

Test whether the number of T-shirts sold by all of the vendors changed significantly, using a 95% confidence level.

Answer: Since $\alpha = .05$ and df $= 6 - 1 = 5$, $t_{\text{crit}} = 2.571$ (from Appendix 2).

$$\bar{d} = \frac{\sum d}{n} = \frac{-1}{6} = -0.1667$$

$$s_d = \sqrt{\frac{\sum d^2 - \frac{(\sum d)^2}{n}}{n-1}} = \sqrt{\frac{997 - \frac{(1)^2}{6}}{6-1}}$$

$$= \sqrt{996.833/5} = \sqrt{199.37} = 14.12$$

$$t_{\text{ob}} = \frac{\bar{d}}{s_d/\sqrt{n}} = \frac{-0.1667}{14.12/\sqrt{6}}$$

$$= \frac{-0.1667}{5.764} = -0.0289$$

Since $t_{\text{crit}} > |t_{\text{ob}}|$, we fail to reject the null hypothesis and conclude that the difference is not statistically significant.

12. During the semester, I have secretly been keeping track of how many classes each of you have missed. I have also (not so secretly) been keeping track of your grades in the class. The data for nine of your classmates, with names removed to protect the guilty, appears below:

Subject Number	Number of Classes Skipped (x)	Average (y)
1	0	86
2	1	92
3	3	76
4	2	84
5	6	45
6	3	82
7	0	94
8	7	54
9	1	83

Assume the causal model *Skipping class* $\rightarrow$ *lower grades*. Estimate: the slope b and intercept a of the regression line, the correlation coefficient r, and the coefficient of determination r^2. Also, test the null hypothesis that $r = 0$ with a confidence level of 99%. What does this say about our model (i.e. does it indicate it is accurate)?

HINT: It may be helpful to figure out $\sum X$, $\sum Y$, $\sum X^2$, $\sum Y^2$ and $\sum XY$ before you get started.

Answer: Since $\alpha = .01$ and df $= 9 - 2 = 7$, $t_{\text{crit}} = 3.499$ (from Appendix 2).

$$\sum X = 23 \qquad \sum Y = 696$$
$$\sum X^2 = 109 \qquad \sum Y^2 = 56082$$
$$\sum XY = 1465$$

$$b = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{1465 - (23)(696)/9}{109 - 23^2/9}$$
$$= \frac{1465 - 16008/9}{109 - 529/9} = \frac{1465 - 1778.67}{109 - 58.78} = \frac{-313.67}{50.22} = -6.246$$
$$a = \bar{y} - b\bar{x} = (696/9) - (-6.246)(23/9) = 77.33 - (-6.246)(2.556) = 77.33 + 15.96 = 93.29$$

$$
\begin{aligned}
r &= \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{\sum X^2 - (\sum X)^2/n}\,\sqrt{\sum Y^2 - (\sum Y)^2/n}} = \frac{-313.67}{\sqrt{50.22}\,\sqrt{56082 - 696^2/9}} \\
&= \frac{-313.67}{\sqrt{50.22}\,\sqrt{56082 - 484416/9}} = \frac{-313.67}{\sqrt{50.22}\,\sqrt{56082 - 53824}} \\
&= \frac{-313.67}{\sqrt{50.22}\,\sqrt{2258}} = \frac{-313.67}{(7.09)(47.52)} = \frac{-313.67}{336.91} = -0.9314 \\
r^2 &= (r)^2 = (-0.9314)^2 = 0.8676. \\
t_{\text{ob}} &= r\sqrt{\frac{n-2}{1-r^2}} = -0.9314\sqrt{\frac{9-2}{1-0.8676}} = -0.9314\sqrt{7/0.1324} \\
&= -0.9314\sqrt{52.88} = -0.9314(7.272) = -6.7732
\end{aligned}
$$

Since $t_{\text{crit}} < |t_{\text{ob}}|$, we reject the null hypothesis and conclude that the correlation between x and y is statistically significant. This suggests our model is plausible.