5.4

(a) A 95% confidence interval can be calculated as follows:

$$
\begin{aligned}
\bar{x}_{diff} \pm z^{\star} \frac{s_{diff}}{\sqrt{n}} &= -0.545 \pm 1.96 \times \frac{8.887}{\sqrt{200}} \\
&= -0.545 \pm 1.96 \times 0.6284 \\
&= -0.545 \pm 1.231 \\
&= (-1.78, 0.69)
\end{aligned}
$$

(b) We are 95% confident that on the reading test students score, on average, 1.78 points lower to 0.69 points higher than they do on the writing test.

(c) No, since 0 is included in the interval.

5.14 The 95% confidence interval can be calculated as follows:

$$\begin{aligned}(\bar{x}_f - \bar{x}_m) \pm z^\star \sqrt{\frac{s_f^2}{n_f} + \frac{s_m^2}{n_m}} &= (31 - 16) \pm 1.96 \times \sqrt{\frac{31^2}{487} + \frac{21^2}{312}} \\ &= 15 \pm 1.96 \times 1.84 \\ &= 15 \pm 3.6 \\ &= (11.4, 18.6)\end{aligned}$$

We are 95% confident that Chinese females spend on average 11.4 to 18.6 hours per week more than males taking care of their children under age 6. Since this interval doesn't contain 0, it indicates a significant difference between the average number of hours Chinese males and females spend taking care of their children under age 6.

5.20

(a) This may not be reasonable since the sample may not be random. The drivers who volunteer to submit their gas mileage on fueleconomy.gov might be those that are getting much lower or much higher than the gas mileage estimated by the EPA.

(b) The hypotheses are: $H_0 : \mu = 50, H_A : \mu \neq 50$.
Before calculating the test statistic, we should evaluate the conditions for the test:

1. Independence: Independence: Our sample is a convenience sample, which is a red flag regarding the independence of observations (even when limiting our population to be those who participate on the fueleconomy.gov website). When reporting these results to others, we should volunteer this information and note that our results rely on the assumption that the observations are independent.
2. Sample size: The sample size is less than 30, therefore we will use the t distribution.
3. Skew: The distribution is approximately symmetric and there is no evidence that it is not nearly normal, though checking this conditions is difficult for such a small sample.

The test statistic and the p-value can be calculated as follows:

$$T_{df} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{53.3 - 50}{\frac{5.2}{\sqrt{14}}} = \frac{3.3}{1.39} = 2.37$$
$$df = n - 1 = 14 - 1 = 13$$
$$0.02 < p - value < 0.05$$

Since p-value < 0.05, reject H_0. The data provide strong evidence against the EPA claim of 50 MPG.

(c) Given that $T^\star = 2.16$, a 95% confidence interval can be calculated as follows:

$$\begin{aligned} \bar{x} \pm t^\star_{df} \frac{s}{\sqrt{n}} &= 53.3 \pm 2.16 \frac{5.2}{\sqrt{14}} \\ &= 53.3 \pm 2.16 \times 1.39 \\ &= 53.3 \pm 3 \\ &= (50.3, 56.3) \end{aligned}$$

We are 95% confident that a 2012 Prius gets on average 50.3 to 56.3 MPG.

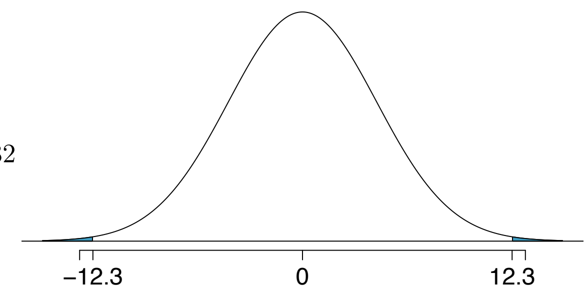
5.26 The hypotheses are: $H_0 : \mu_{0.99} = \mu_1$ and $H_A : \mu_{0.99} \neq \mu_1$. The conditions that need to be satisfied for the sampling distribution of $(\bar{x}_{0.99} - \bar{x}_1)$ to be nearly normal and the estimate of the standard error to be sufficiently accurate are:

1. Independence: Both samples are random and represent less than 10% of their respective populations. Also, we have no reason to think that the 0.99 carats are not independent of the 1 carat diamonds since they are both sampled randomly.
2. Sample size: Sample size is less than 30, therefore we use a t-test.
3. Skew: The distributions are not extremely skewed.

The test statistic and the p-value are calculated as follows:

$$\begin{aligned} T &= \frac{(\bar{x}_{0.99} - \bar{x}_1) - (\mu_{0.99} - \mu_1)}{\sqrt{\frac{s^2_{0.99}}{n_{0.99}} + \frac{s^2_1}{n_1}}} \\ &= \frac{(44.51 - 56.81) - 0}{\sqrt{\frac{13.32^2}{23} + \frac{16.13^2}{23}}} = \frac{-12.3}{4.36} = -2.82 \\ df &= 23 - 1 = 22 \end{aligned}$$

$$p - value = 0.01$$



Since p-value < 0.05, reject H_0. The data provide convincing evidence that the average standardized price of 0.99 carats and 1 carat diamonds are different.

5.28 The 95% confidence interval can be calculated as follows:

$$t^\star_{df} = t^\star_{23-1=22} = 2.07$$

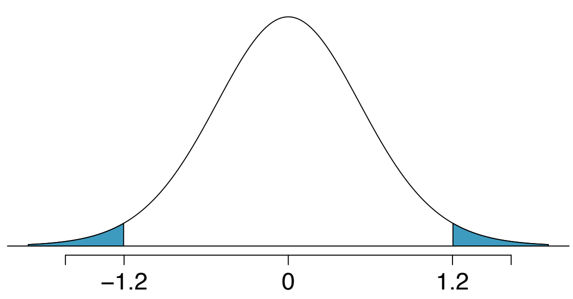
$$\begin{aligned} (\bar{x}_{0.99} - \bar{x}_1) \pm t^\star_{df}\sqrt{\frac{s^2_{0.99}}{n_{0.99}} + \frac{s^2_1}{n_1}} &= (56.81 - 44.51) \pm 2.07\sqrt{\frac{s^2_{0.99}}{n_{0.99}} + \frac{s^2_1}{n_1}} \\ &= (44.51 - 56.81) \pm 2.07\sqrt{\frac{13.32^2}{23} + \frac{16.13^2}{23}} \\ &= -12.3 \pm 2.07 \times 4.36 \\ &= -12.3 \pm 9.03 \\ &= (-21.33, -3.27) \end{aligned}$$

We are 95% confident that the average standardized price of a 0.99 carat diamond is \$3.27 to \$21.33 lower than the average standardized price of a 1 carat diamond.

5.34 The hypotheses are $H_0 : \mu_T = \mu_C$ and $H_A : \mu_T \neq \mu_C$.

We are told to assume that conditions for inference are satisfied, population standard deviation is unknown and sample sizes are small, hence we calculate a T score.

$$
\begin{aligned}
T &= \frac{(\bar{x}_T - \bar{x}_C) - (\mu_T - \mu_C)}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_C^2}{n_C}}} \\
&= \frac{(4.9 - 6.1) - 0}{\sqrt{\frac{1.8^2}{22} + \frac{1.8^2}{22}}} = \frac{-1.2}{0.543} = -2.21 \\
df &= min(n_1 - 1, n_2 - 1) = min(22 - 1, 22 - 1) = 21 \\
p - value &= P(|t_{21}| > 2.21) \\
0.02 < p &- value < 0.05
\end{aligned}
$$



Since p-value < 0.05, we reject H_0. The data provide convincing evidence that the average number of food items recalled by the patients in the treatment and control groups are different.

6.2

(a) TRUE. The success-failure condition is not satisfied

$$np = 20 \times 0.77 = 15.4 \; and \; n(1-p) = 20 \times 0.23 = 4.6,$$

therefore we know that the distribution of $\hat{p}$ is not approximately normal. In most samples we would expect $\hat{p}$ to be close to 0.77, the true population proportion. While $\hat{p}$ can be as low as 0 (though we would expect this to happen very rarely), it can only go as high as 1. Therefore, since 0.77 is closer to 1, the distribution would probably take on a left skewed shape. Plotting the sampling distribution would confirm this suspicion.

(b) FALSE. Unlike with means, for the sampling distribution of proportions to be approximately normal, we need to have at least 10 successes and 10 failures in our sample. We do not use $n \geq 30$ as a condition to check for the normality of the distribution of $\hat{p}$.

(c) FALSE. Standard error of $\hat{p}$ in samples with $n = 60$ can be calculated as:

$$SE_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.77 \times 0.23}{60}} = 0.0543$$

A $\hat{p}$ of 0.85 is only $Z = \frac{0.85-0.77}{0.0543} = 1.47$ standard errors away from the mean, which would not be considered unusual.

(d) TRUE. Standard error of $\hat{p}$ in samples with $n = 180$ can be calculated as:

$$SE_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.77 \times 0.23}{180}} = 0.0314$$

A $\hat{p}$ of 0.85 is $Z = \frac{0.85-0.77}{0.0314} = 2.54$ standard errors away from the mean, which might be considered unusual.