STA 032 Homework 7

Hardy Jones
999397426
Professor Melcon
Winter 2015

§ 5.2 1 We have $n = 70$ independent Bernoulli trials with $X = 28$ successes.

(1) $\frac{28}{70} = 0.4 = 40\%$ of the sampled automobiles had emission levels that exceeded the standard.

(2) Taking $p = 0.4$ we have a $X \sim Bin(70, 0.4)$.
We take $\tilde{n} = 70 + 4 = 74$, $\tilde{p} = \frac{28+2}{74} = 0.\overline{405}$.
A 95% confidence interval can now be found.

$$\begin{aligned}
\tilde{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} &\approx 0.405 \pm z_{0.025} \sqrt{\frac{0.405(1-0.405)}{74}} \\
&\approx 0.405 \pm 1.96\sqrt{0.00326} \\
&\approx 0.405 \pm 0.112
\end{aligned}$$

So the interval is (0.293, 0.517).

(3) Using the values calculated before, a 98% confidence interval can be found.

$$\begin{aligned}
\tilde{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} &\approx 0.405 \pm z_{0.01} \sqrt{\frac{0.405(1-0.405)}{74}} \\
&\approx 0.405 \pm 2.33\sqrt{0.00326} \\
&\approx 0.405 \pm 0.133
\end{aligned}$$

So the interval is (0.272, 0.538).

(4) Since we're going to be solving more than one of these similar questions, let's find a closed form to calculate this easily.
For each problem we need to know $\alpha, \tilde{p}$ and the range r.
We want to solve the following equation for n:

$$
\begin{aligned}
r &= z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \\
&= z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \\
\frac{r}{z_{\frac{\alpha}{2}}} &= \sqrt{\frac{\tilde{p}(1-\tilde{p})}{n+4}} \\
\left(\frac{r}{z_{\frac{\alpha}{2}}}\right)^2 &= \frac{\tilde{p}(1-\tilde{p})}{n+4} \\
n+4 &= \frac{\tilde{p}(1-\tilde{p})}{\left(\frac{r}{z_{\frac{\alpha}{2}}}\right)^2} \\
n &= \frac{\tilde{p}(1-\tilde{p})}{\left(\frac{r}{z_{\frac{\alpha}{2}}}\right)^2} - 4 \\
n &= z_{\frac{\alpha}{2}}^2\frac{\tilde{p}(1-\tilde{p})}{r^2} - 4 \qquad (1)
\end{aligned}
$$

Now with equation 1 we can solve with plug and chug.

$$n = 1.96^2\frac{0.405(1-0.405)}{0.10^2} - 4 = 88.573$$

So 89 samples are needed for the proportion to exceed the standard to within $\pm$ 0.10 with 95% confidence.

(5) Using equation 1 we can solve with plug and chug.

$$n = 2.33^2\frac{0.405(1-0.405)}{0.10^2} - 4 = 126.823$$

So 127 samples are needed for the proportion to exceed the standard to within $\pm$ 0.10 with 98% confidence.

(6) We would do well to also find a closed form for this question.
We need the $\tilde{n}, \tilde{p}$ and the upper bound u.
Then we can solve the following equation for z_α:

$$
\begin{aligned}
u &= \tilde{p} + z_\alpha \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \\
u - \tilde{p} &= z_\alpha \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} \\
\frac{u-\tilde{p}}{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}} &= z_\alpha \\
z_\alpha &= \frac{u-\tilde{p}}{\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}}} \\
z_\alpha &= (u-\tilde{p})\sqrt{\frac{\tilde{n}}{\tilde{p}(1-\tilde{p})}} \qquad (2)
\end{aligned}
$$

Now, using equation 2 we can solve for an upper bound of 0.50:

$$
\begin{aligned}
z_\alpha &= (0.50 - 0.405)\sqrt{\frac{74}{0.405(1-0.405)}} \\
&= 1.66
\end{aligned}
$$

So the z score corresponds to $0.9515 = 95.15\%$.
Thus we can say with 95.15% confidence that less than half of the vehicles in the state exceed the standard.

4 We have $n = 444$ independent Bernoulli trials with $X = 170$ successes.
$\frac{170}{444} = 0.382\overline{882}$ of the sampled smokers used the patch.
So we have $X \sim Bin(444, 0.383)$.
We also calculate $\tilde{n} = 444 + 4 = 448$ and $\tilde{p} = \frac{170+2}{448} = 0.384$

(1) A 95% confidence interval can be found:

$$
\begin{aligned}
\tilde{p} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} &= 0.384 \pm 1.96\sqrt{\frac{0.384(1-0.384)}{448}} \\
&= 0.384 \pm 0.0450
\end{aligned}
$$

So the interval is (0.339, 0.429).

(2) A 98% confidence interval can be found:

$$\tilde{p} \pm z_{\frac{\alpha}{2}}\sqrt{\frac{\tilde{p}(1-\tilde{p})}{\tilde{n}}} = 0.384 \pm 2.58\sqrt{\frac{0.384(1-0.384)}{448}}$$
$$= 0.384 \pm 0.0592$$

So the interval is (0.325, 0.443).

(3) Using equation 2 we can solve for an upper bound of 0.40:

$$z_\alpha = (0.40 - 0.384)\sqrt{\frac{448}{0.384(1-0.384)}}$$
$$= 0.70$$

So the z score corresponds to 0.7580 = 75.80%.
Thus we can say with 75.80% confidence that proportion is less than 0.40.

(4) Using equation 1 we can solve with plug and chug.

$$n = 1.96^2\frac{0.384(1-0.384)}{0.03^2} - 4 = 1005.675$$

So 1006 samples are needed for a 95% confidence to specify the proportion within ± 0.03.

(5) Using equation 1 we can solve with plug and chug.

$$n = 2.58^2\frac{0.384(1-0.384)}{0.03^2} - 4 = 1749.479$$

So 1750 samples are needed for a 99% confidence to specify the proportion within ± 0.03.

§ 5.3 10 We have $n = 15, \overline{X} = 13, s = 2$.
We compute

$$13 \pm 2.977\frac{2}{\sqrt{15}} = (11.463, 14.537)$$

So we can say with 99% confidence that the mean track length is in the interval (11.463µm, 14.537µm).

11 We have $n = 6, \overline{X} = 2.03, s = 0.090$.
We compute

$$2.03 \pm 2.015\frac{0.090}{\sqrt{6}} = (1.956, 2.104)$$

So we can say with 90% confidence that the mean deflection caused by a 160kN load is in the interval (1.956mm, 2.104mm).

§ 5.4 3 We have $n_X = 1559, \overline{X} = 30.4, \sigma_X = 0.6$
and $n_Y = 1924, \overline{Y} = 31.1, \sigma_Y = 0.2$
We can find a 99% confidence interval as follows:

$$\overline{X} - \overline{Y} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} = 30.4 - 31.1 \pm 2.576\sqrt{\frac{0.6^2}{1559} + \frac{0.2}{1924}} = (-0.741, -0.659)$$

So we can say with 99% confidence that men have between -0.741 and -0.659 lower BMI than women.

4 We have $n_X = 296, \overline{X} = 54.1, \sigma_X = 4.4$
and $n_Y = 296, \overline{Y} = 72.7, \sigma_Y = 4.7$
We can find a 95% confidence interval as follows:

$$\overline{X} - \overline{Y} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}} = 54.1 - 72.7 \pm 1.960\sqrt{\frac{4.4^2}{296} + \frac{4.7}{296}} = (-19.333, -14.867)$$

So we can say with 95% confidence that 27 week-old hens lay eggs that weight between -19.333g and -14.867g lower than 59 week-old hens.

§ 5.5 4 We have $X = 43, n_X = 50, p_X = \frac{43}{50} = 0.86, \tilde{n}_X = 50 + 2 = 52, \tilde{p}_X = \frac{43+1}{50+2} = 0.846$
and $Y = 25, n_Y = 40, p_Y = \frac{25}{40} = 0.625, \tilde{n}_Y = 40 + 2 = 42, \tilde{p}_Y = \frac{25+1}{40+2} = 0.619$
We can find a 99% confidence interval as follows:

$$\begin{aligned}
&\tilde{p}_X - \tilde{p}_Y \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}_X(1-\tilde{p}_X)}{\tilde{n}_X} + \frac{\tilde{p}_Y(1-\tilde{p}_Y)}{\tilde{n}_Y}} \\
&= 0.846 - 0.619 \pm 2.576\sqrt{\frac{0.846(1-0.846)}{52} + \frac{0.619(1-0.619)}{42}} \\
&= 0.227 \pm 0.232
\end{aligned}$$

So we can say with 99% confidence that 22.7% ± 23.2% students felt more confident learning from a GUI.

5 We have $X = 8, n_X = 12, p_X = \frac{8}{12} = 0.\overline{6}, \tilde{n}_X = 12 + 2 = 14, \tilde{p}_X = \frac{8+1}{12+2} = 0.643$
and $Y = 5, n_Y = 15, p_Y = \frac{5}{15} = 0.\overline{3}, \tilde{n}_Y = 15 + 2 = 17, \tilde{p}_Y = \frac{5+1}{15+2} = 0.353$
We can find a 95% confidence interval as follows:

$$\begin{aligned}
&\tilde{p}_X - \tilde{p}_Y \pm z_{\frac{\alpha}{2}} \sqrt{\frac{\tilde{p}_X(1-\tilde{p}_X)}{\tilde{n}_X} + \frac{\tilde{p}_Y(1-\tilde{p}_Y)}{\tilde{n}_Y}} \\
&= 0.643 - 0.353 \pm 1.960\sqrt{\frac{0.643(1-0.643)}{14} + \frac{0.353(1-0.353)}{17}} \\
&= 0.290 \pm 0.339
\end{aligned}$$

So we can say with 95% confidence that 29.0% $\pm$ 33.9% more small cars were totaled than large cars.

§ 5.6 8 We have $n_X = 10, \overline{X} = 3.4, s_X = 0.6$
and $n_Y = 15, \overline{Y} = 7.9, s_Y = 0.6$.
We calculate ν as follows:

$$\nu = \left\lfloor \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{s_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{s_Y^2}{n_Y}\right)^2}{n_Y - 1}} \right\rfloor = \left\lfloor \frac{\left(\frac{0.6^2}{10} + \frac{0.6^2}{15}\right)^2}{\frac{\left(\frac{0.6^2}{10}\right)^2}{10-1} + \frac{\left(\frac{0.6^2}{15}\right)^2}{15-1}} \right\rfloor = \lfloor 19.\overline{4} \rfloor = 19$$

We can find a 98% confidence interval as follows:

$$\overline{X} - \overline{Y} \pm t_{\nu, \frac{\alpha}{2}} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}} = 3.4 - 7.9 \pm 2.539 \sqrt{\frac{0.6^2}{10} + \frac{0.6^2}{15}} = -4.5 \pm 0.622$$

So we can say with 98% confidence that patients with higher blood pressure had lower levels of insulin with a mean of $4.5 \pm 0.622\ 10^{-5}\ \text{min}^{-1}\text{pmol}^{-1}$ than patients with normal blood pressure.

9 We have $n_X = 24, \overline{X} = 4.8, s_X = 1.9$
and $n_Y = 24, \overline{Y} = 2.8, s_Y = 1.0$.
We calculate ν as follows:

$$\nu = \left\lfloor \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{s_X^2}{n_X}\right)^2}{n_X - 1} + \frac{\left(\frac{s_Y^2}{n_Y}\right)^2}{n_Y - 1}} \right\rfloor = \left\lfloor \frac{\left(\frac{1.9^2}{24} + \frac{1.0^2}{24}\right)^2}{\frac{\left(\frac{1.9^2}{24}\right)^2}{24-1} + \frac{\left(\frac{1.0^2}{24}\right)^2}{24-1}} \right\rfloor = \lfloor 34.834 \rfloor = 34$$

Using 35 as the t value (as 34 is not in the table), we can find a 95% confidence interval as follows:

$$\overline{X} - \overline{Y} \pm t_{\nu, \frac{\alpha}{2}} \sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}} = 4.8 - 2.8 \pm 2.030 \sqrt{\frac{1.9^2}{24} + \frac{1.0^2}{24}} = 2 \pm 0.890$$

So we can say with 95% confidence that undergraduate students read a passage slower with a mean of 2 $\pm$ 0.89 seconds than graduate students.