Stat 3011 Midterm 2 (Class Part)

Problem 1

This is a *one sample* problem about *proportions*. The parameter is p, the estimator is $\hat{p}$ and the standard error is

$$\text{se}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

In this problem

$$\begin{aligned} \hat{p} &= 0.31 \\ n &= 200 \end{aligned}$$

The standard error is

$$\begin{aligned} \text{se}(\hat{p}) &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \\ &= \sqrt{\frac{0.31 \times 0.69}{200}} \\ &= 0.03270321 \end{aligned}$$

So two standard errors is $2 \times 0.03270321 = 0.06540642$, and the two standard error interval is 0.31 ± 0.065 or $(0.245, 0.375)$.

Problem 2

This is a *two sample* problem about *means*. The parameter is $\mu_1 - \mu_2$, the estimator is $\bar{x}_1 - \bar{x}_2$ and the standard error is

$$\text{se}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

In this problem

$$\begin{aligned} \bar{x}_1 &= 1150 \\ s_1 &= 1020 \\ n_1 &= 100 \\ \bar{x}_2 &= 1050 \\ s_2 &= 1030 \\ n_2 &= 200 \end{aligned}$$

So the estimator is

$$\bar{x}_1 - \bar{x}_2 = 1150 - 1050 = 100$$

and the standard error is

$$
\begin{aligned}
\operatorname{se}(\bar{x}_1 - \bar{x}_2) &= \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \\
&= \sqrt{\frac{1020^2}{100} + \frac{1030^2}{200}} \\
&= 125.3336
\end{aligned}
$$

(a) The problem asks for an approximate, large sample confidence interval so a z critical values will do. The z critical value for 95% confidence is 1.96. Hence the confidence interval is

$$\bar{x}_1 - \bar{x}_2 \pm 1.96 \operatorname{se}(\bar{x}_1 - \bar{x}_2)$$

and

$$1.96 \operatorname{se}(\bar{x}_1 - \bar{x}_2) = 1.96 \times 125.3336 = 245.7$$

And the interval is 100 ± 245.7 or $(-145.7, 345.7)$.

(b) The z critical value for 90% confidence is 1.645. Hence the confidence interval is

$$\bar{x}_1 - \bar{x}_2 \pm 1.645 \operatorname{se}(\bar{x}_1 - \bar{x}_2)$$

and

$$1.645 \operatorname{se}(\bar{x}_1 - \bar{x}_2) = 1.645 \times 125.3336 = 206.2$$

And the interval is 100 ± 206.2 or $(-106.2, 306.2)$.

Problem 3

This is a question about sample size calculation (Sections 8.6.1 and 8.6.2 in Wild and Seber). It is clearly a problem about *proportions*, so the appropriate formula is

$$n = \left(\frac{z}{m}\right)^2 \times p^*(1 - p^*)$$

where

- z is the critical value. For 95% confidence $z = 1.96$.
- m is the desired margin of error. Here $m = 0.02$.
- p^* is a guess about the unknown true propulation proportion p. When one wants a sample size that works for any p, which is what we want here, use $p^* = 0.5$.

So

$$
\begin{aligned}
n &= \left(\frac{z}{m}\right)^2 \times p^*(1-p^*) \\
&= \left(\frac{1.96}{0.02}\right)^2 \times 0.5 \times 0.5 \\
&= 2401
\end{aligned}
$$

Problem 4

(a) A "gimmie." The 3% margin of error reported for the poll.

(b) This is the adjustment for subgroup size (square root law). Multiply the margin of error by the square root of the ratio of the whole sample size to the subgroup size, in this case $\sqrt{16} = 4$. So $4 \times 3\% = 12\%$.

(c) This is "case (c)" in Wild and Seber (**one** sample, **different** questions). Multiply the margin of error by 2. So $2 \times 3\% = 6\%$.

(d) This is a "case (c)" difference of proportions involving a subgroup. Apply both adjustments, the subgroup adjustment from part (b) and the difference adjustment from part (c). So $4 \times 2 \times 3\% = 24\%$.