Chapter 22 – Inference for a Population Proportion

22.1 (a) The population is surgical patients. p is the proportion of all surgical patients who will test positive for *Staphylococcus aureus*. **(b)** $\hat{p} = \frac{1251}{6771} = 0.185$, or 18.5%.

22.3 (a) Because np = 0.90(1500) = 1350 and $n(1 - p)$ = 0.10(1500) = 150, and both are at least 10, the sampling distribution of $\hat{p}$ is approximately Normal, with mean p = 0.90 and standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.90(1-.90)}{1500}}$ = 0.0077. **(b)** If n = 6000, the sampling distribution of $\hat{p}$ is approximately Normal, with mean p = 0.90 and standard deviation $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.90(1-0.90)}{6000}}$ = 0.0039. Notice that quadrupling the sample size (from 1500 to 6000) results in halving the standard deviation of $\hat{p}$ (0.0039 is one-half of 0.0077, at least to within rounding).

22.5 (a) The survey excludes residents of the northern territories, as well as those who have no phones or have only cell phone service. **(b)** $\hat{p} = \frac{1288}{1505}$ = 0.8558, so SE = $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ = 0.009055, and the 95% confidence interval is 0.8558 ± (1.96)(0.009055) = 0.8381 to 0.8735, or 83.8% to 87.4%.

22.7 (a) We have $\hat{p} = \frac{42}{165} = 0.255$, so the margin of error is $1.96\sqrt{\frac{0.255(1-0.255)}{165}}$ = 0.0665. **(b)** For a ±3% margin of error, we'll need more than four times this sample size, because 3% is less than half the original margin of error. The actual number needed in the sample (using the original value of $\hat{p}$ as p^*) is $n = \left(\frac{1.96}{0.03}\right)^2 (0.255)(1-0.255) = 810.898$, so we need at least 811 visitors over age 65.

22.9 STATE: We wonder if the proportion of times the Belgian euro coin spins heads is the same as the proportion of times it spins tails. SOLVE: Let p be the proportion of times a spun Belgian euro coin lands heads. We test $H_0: p = 0.50$ versus $H_a: p \neq 0.50$. Because the sample consists of 250 trials, we expect 125 "successes" (heads) and 125 "failures" (tails); both of these are at least 10, and we assume the sample represents an SRS of all possible coin spins, so the conditions are met. $\hat{p}$ = $\frac{140}{250}$ = 0.56, and SE = $\sqrt{\frac{p_0(1-p_0)}{n}}$ = $\sqrt{\frac{0.50(1-0.50)}{250}}$ = 0.0316. $z = \frac{\hat{p}-p_0}{\text{SE}}$ = $\frac{0.56-0.50}{0.0316}$ = 1.90, and P = 0.0574. CONCLUDE: There is some evidence that the proportion of times a Belgian euro coin spins heads is not 0.50; the P-value is close to 0.05, but not less than 0.05. Perhaps more spins would be conclusive.

22.15 (b) The sampling distribution of $\hat{p}$ has mean p = 0.60.

22.16 (b) The standard deviation of $\hat{p}$ is $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.60(1-0.60)}{117}}$ = 0.0453.

22.17 (b) The 90% confidence interval is 0.80 ±1.645 $\sqrt{\frac{0.80(1-0.80)}{4500}}$.

22.18 (a) Less confidence means a smaller margin of error.

22.19 (c) $n = \left(\frac{z^*}{m}\right)^2 p^*(1-p^*) = \left(\frac{2.576}{0.02}\right)^2 (0.5)(0.5)$ = 4147.36, so we would need at least 4148.

22.20 (b) With $\hat{p} = \frac{53}{100}$ = 0.53, the margin of error is 1.96$\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ = 1.96 $\sqrt{\frac{0.53(1-0.53)}{100}}$ = 0.098.

22.21 (a) Sources of bias are not accounted for in a margin of error.

22.22 (a) The alternative hypothesis expresses the idea "more will finish the maze faster with the loud noise."

22.23 (c) The P-value is 0.0057.

22.24 (a) This is what 95% confidence means.

22.25 (a) The survey excludes those who have no phones or have only cell phone service. **(b)** Note that we have 848 "Yes" answers and 162 "no" answers; both of these are at least 15. With the sample proportion $\hat{p} = \frac{848}{1010}$ = 0.8396, the large-sample 95% confidence interval is $0.8396 \pm 1.96\sqrt{\frac{0.8396(1-0.8396)}{1010}}$ = 0.8170 to 0.8622.

22.35 PLAN: With p representing the proportion of songs downloaded by Rina, we test $H_0: p = 0.50$ versus $H_a: p \neq 0.50$. The test is two-sided because we wonder if the proportion downloaded by Rina differs from that downloaded by Ed, which would mean that p differs from 0.5. SOLVE: We assume that the 50 songs sampled are an SRS. With 50 songs sampled, we expect 50(0.50) = 25 successes and 50(0.50) = 25 failures (which are both at least 10), so conditions for use of the large-sample z test are satisfied. We have $\hat{p} = \frac{34}{50}$ = 0.68, so $z = \frac{0.68 - 0.50}{\sqrt{\frac{0.50(1-0.50)}{50}}}$ = 2.55 and P = 0.0108. CONCLUDE: There is strong evidence that the proportion of songs downloaded by Rina differs from 0.50. In fact, it seems that Rina downloaded more than Ed. **(b)** The conditions for a large-sample confidence interval are met because 34 of the sample's 50 songs were downloaded by Rina and 16 by Ed; both of these are larger than 15. The 95% confidence interval for the proportion downloaded by Rina is $0.68 \pm 1.96\sqrt{\frac{0.68(1-0.68)}{50}}$ = 0.5507 to 0.8093. At 95% confidence, Rina has downloaded between about 55% and 81% of the songs on their player.