

HW 14 [PG]: **You don't need to turn in this set of homeworks**

Chapter 14: 5, 8, 11

Chapter 15: 8, 12

Ch 14, 5: a. The exact binomial probability that 4 or fewer of the infants weigh at most 2500 grams is

$$\begin{aligned} P(X \leq 4) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= \sum_{x=0}^4 \binom{40}{x} 0.15^x 0.85^{40-x} \\ &= 0.263. \end{aligned}$$

b. Since $np = 40 * 0.15 = 6$ and $n(1 - p) = 40 * 0.85 = 34$ are both greater than 5, we can use the normal approximation to the binomial distribution. Applying the continuity correction, we find that

$$z = \frac{X - np + 0.5}{\sqrt{np(1 - p)}} = \frac{4 - 40 \times 0.15 + 0.5}{\sqrt{40(0.15)(0.85)}} = -0.66.$$

The area under the standard normal curve that lies to the left of $z = -0.66$ is 0.255; this is the estimated probability that at most 4 of the newborns weight at most 2500 grams.

c. The normal approximation provides a fairly good estimate of the exact probability.

Ch 14, 8: a.

$$\hat{p} = \frac{473}{488} = 0.969.$$

b. Since $n\hat{p} = 488 \times 0.969 = 473$ and $n(1 - \hat{p}) = 15$, the sample size is large enough to justify the use of the normal approximation. Therefore, an approximate 95% confidence interval for p is

$$\left(0.969 - 1.96\sqrt{\frac{0.969(1 - 0.969)}{488}}, 0.969 + 1.96\sqrt{\frac{0.969(1 - 0.969)}{488}} \right) = (0.954, 0.984).$$

c. If 100 random samples of size 488 were drawn from this population and 100 different 95% confidence intervals generated, approximately 95 would contain the true population proportion p and 5 would not.

d. An approximate 90% confidence interval for p is

$$\left(0.969 - 1.645\sqrt{\frac{0.969(1 - 0.969)}{488}}, 0.969 + 1.645\sqrt{\frac{0.969(1 - 0.969)}{488}} \right) = (0.956, 0.982).$$

e. The 90% confidence interval is shorter than the corresponding 95% interval.

Ch 14, 11: a. For individuals assigned to the prepaid plan, the estimated proportion of patients who visited a community crisis center is

$$\hat{p}_1 = \frac{13}{311} = 0.042.$$

Among those receiving traditional Medicaid,

$$\hat{p}_2 = \frac{22}{310} = 0.071.$$

b.

$$H_0 : p_1 = p_2, \quad H_A : p_1 \neq p_2.$$

The pooled estimate of the common proportion is

$$\hat{p} = \frac{13 + 22}{311 + 310} = 0.056.$$

The test statistic is

$$z = \frac{0.042 - 0.071 - 0}{\sqrt{0.056(1 - 0.056)[(1/311) + (1/310)]}} = -1.57.$$

. The p-value $p = 2 \times 0.058 = 0.116$. We are unable to reject the null hypothesis at the 0.10 level of significance.

c. There is insufficient evidence to conclude that the proportions of patients visiting a community crisis center are not identical to those on the prepaid medical plan and those receiving traditional Medicaid.

Ch15, 8: To test the null hypothesis that there is no association between driving while drinking and calendar year, we use the chi-square test. We first calculate the table of expected counts.

Drove while drinking	1983	1987	Total
Yes	1116.3	1124.7	2241
No	1520.7	1523.3	3053
Total	2637	2657	5294

The test statistic is

$$\begin{aligned} \chi^2 &= \sum_{i=1}^4 \frac{(|O_i - E_i| - 0.5)^2}{E_i} \\ &= \frac{133.2^2}{1116.3} + \frac{133.2^2}{1124.7} + \frac{133.2^2}{1520.7} + \frac{133.2^2}{1523.3} \\ &= 15.89 + 15.78 + 11.67 + 11.58 \\ &= 54.92. \end{aligned}$$

The df= 1. The p-value $p < 0.001$. We reject the null hypothesis at the 0.05 level of significance.

b. There does appear to be an association between driving while drinking and calendar year. A small proportion of college students were driving while drinking in 1987 ($991/2657 = 37\%$), after the legal drinking age was raised, than were doing so in 1983 ($1250/2637 = 47\%$).

c. To begin, the pooled estimate of the common proportion is

$$\hat{p} = \frac{1250 + 991}{5294} = 0.423.$$

The test statistic

$$z = \frac{0.474 - 0.373}{\sqrt{0.423(1 - 0.423)\left[\frac{1}{2637} + \frac{1}{2657}\right]}} = 7.44.$$

The p-value $p < 0.001$. We reject the null hypothesis at the 0.05 level of significance. Therefore, we do reach the same conclusion as we did with the chi-square test.

d. The estimated difference in proportions is

$$\hat{p}_1 - \hat{p}_2 = 0.474 - 0.373 = 0.101.$$

Using the normal approximation, an approximate 95% confidence interval for the true difference $p_1 - p_2$ is

$$0.101 \pm \sqrt{\frac{0.474(1 - 0.474)}{2637} + \frac{0.373(1 - 0.373)}{2657}} = (0.075, 0.127).$$

e. The 95% confidence interval does not cover the value 0. Since we rejected the null hypothesis that the proportions are identical, we would not have expected that it would.

Ch 15, 12: a. To test the null hypothesis that there is no association between time of screening and diagnosis, we use McNemar's test. The test statistic is

$$\chi^2 = \frac{[|r - s| - 1]^2}{r + s} = \frac{[|489 - 403| - 1]^2}{489 + 403} = 8.10.$$

For a chi-square distribution with 1 df, $0.001 < p < 0.01$. Therefore, we reject H_0 at the 0.05 level of significance. The diagnosis is more likely to be present at the first screening and absent at the second than it is to be absent at the first screening and present at the second.

b. In this case, the data are presented as if they were independent rather than paired samples. There are 6650 observations and only 3325 slides. The chi-square test is incorrect for such data.