Are bicycle helmets effective at preventing head injuries? The following data was collected as part of a case-control study published in *The New England Journal of Medicine*. Consider the data to be random samples of all individuals involved in a bicycle accident in a large region over a year-long period of time. One population comprises individuals who wear bicycle helmets and the other comprises those who do not.

	Wearing a helmet	
Head injury	Yes	No
Yes	17	218
No	130	428

(a) What proportion of sample individuals wearing helmets suffered a head injury? What proportion of those in the sample not wearing helmets suffered head injuries? (Please round these proportions to three digits.)

Solution:

The proportion of individuals wearing helmets that suffered a head injury was $17/147 = 0.116$. The proportion of individuals not wearing helmets that suffered a head injury was $218/646 = 0.337$.

(b) Construct a 95% confidence interval for the difference in population proportions of head injuries among individuals who do not wear and who do wear bicycle helmets. Define any population parameters you use in words.

Solution: Let p_{with} and p_{without} be the population proportions of head injuries among individuals in bicycle accidents that do and do not wear helmets. The confidence interval uses the formula

$$(\tilde{p}_1 - \tilde{p}_2) \pm 1.96\sqrt{\frac{\tilde{p}_1(1-\tilde{p}_1)}{n_1+2} + \frac{\tilde{p}_2(1-\tilde{p}_2)}{n_2+2}}$$

where $\tilde{p}_i = (y_i + 1)/(n_i + 2)$ for $i = 1, 2$ are the adjusted sample proportions.

A 95% confidence interval for $p_{\text{with}} - p_{\text{without}}$ is -0.217 ± 0.064, or, $-0.281 < p_{\text{with}} - p_{\text{without}} < -0.153$.

(c) Carry out a chi-square test with a directional alternative that helmets prevent head injuries. State hypotheses, calculate a test statistic, use a table to find a range for the p-value, and interpret the results in the context of the problem.

Solution: Using the above notation, the hypothesis we are testing is

$$H_0\colon\ p_{\text{with}} = p_{\text{without}} \qquad H_A\colon\ p_{\text{with}} < p_{\text{without}}\ .$$

The expected counts are:

$$\begin{bmatrix} 43.6 & 191.4 \\ 103.4 & 454.6 \end{bmatrix}$$

and the chi-square test statistic is 28.3 and there is one degree of freedom. The p-value for the directional test is half the area to the right of the chi-square test statistic. From the table, it is smaller than 0.0001/2, $p < 0.00005$. Using R, it is $5.1e-08$.

There is very strong evidence that individuals in bike accidents wearing helmets were less likely to suffer head injuries than those not wearing helmets.

(d) An investigator carries out Fisher's Exact Test for this problem. The p-value can be thought of as the probability of an event in a random draw of colored balls from a bucket. Fill in the blanks and circle choices to complete this statement correctly.

In Fisher's Exact Test, the p-value for a directional hypothesis test with alternative hypothesis that helmets help to prevent head injuries is the probability of drawing 17 red balls [OR MORE | OR FEWER | EXACTLY] from a bucket with ______________ red balls and ______________ white balls if ______________ balls are drawn from the bucket at random [WITH | WITHOUT] replacement.

Solution: There are two correct solutions.

1. 17 or fewer, 235, 558, 147, without replacement.
2. 17 or fewer, 147, 646, 235, without replacement.