

Instructions

Upload your own solutions (handwritten or typed) to the corresponding Canvas Assignment on or before the due date: 17/12/2021.

Grading scheme

a	b	c	d	e	Total
2	3	8	3	4	20

A survey. A big tech company would like to get an idea of the proportion of their employees who are vaccinated against Covid 19 (at least once). For this reason, they send out a survey.

a. Suppose that the survey is sent to 1000 randomly selected employees. Suppose further that the participation in the study is mandatory and out of the 1000 participants, 730 are fully vaccinated. On their website, the company claims

“We are 95% confident that at least 70% of our employees are vaccinated at least once.”

The claim is based on the large sample 95% confidence interval

$$\left(0.73 - 1.96 \frac{\sqrt{0.73(1-0.73)}}{\sqrt{1000}}, 0.73 + 1.96 \frac{\sqrt{0.73(1-0.73)}}{\sqrt{1000}} \right) = (0.702, 0.758).$$

Is the claim reasonable? Explain.

b. Suppose now that the survey is sent out to all employees but that participation in the survey is voluntary. Out of 422 employees who took the survey, 392 are vaccinated at least once. A corresponding large sample 95% confidence interval for the proportion of vaccinated employees is given by

$$\left(\frac{392}{422} - 1.96 \frac{\sqrt{\frac{392}{422} \left(1 - \frac{392}{422}\right)}}{\sqrt{422}}, \frac{392}{422} + 1.96 \frac{\sqrt{\frac{392}{422} \left(1 - \frac{392}{422}\right)}}{\sqrt{422}} \right) = (0.904, 0.953).$$

Is the following claim justified in this setting?

“We are 95% confident that at least 90% of our employees are vaccinated at least once.”

c. A follow-up study. Suppose that in the setting of part **b**, 100 employees are drawn at random and asked in a follow-up survey about their vaccination status and whether they participated in the initial study. The summary of the results is given in the cross table below.

		Vaccinated		total
		yes	no	
took survey	yes	31	7	38
	no	36	26	62
total		67	33	100

Is there a significant difference between the proportion of vaccinated people that took the survey and those who did not? Conduct the appropriate test in 8 steps at a 5% level of significance.

d. Did you apply a test on independence or a test on homogeneity in part **c**? Explain your reasoning behind your choice.

e. To conclude, let us reflect a bit about parts **a-c**. If you were to design a study based on a survey and both designs (either **a** or **b**) would be an option, which one would you choose and why?