Calculate and interpret confidence intervals for estimating a population proportion.During an election year, we see articles in the newspaper that stateconfidence intervals in terms of proportions or percentages. For example, a poll for a particular candidate running for president might show that the candidate has 40% of the vote within three percentage points (if the sample is large enough). Often, election polls are calculated with 95% confidence, so, the pollsters would be 95% confident that the true proportion of voters who favored the candidate would be between 37% and 43%.Investors in the stock market are interested in the true proportion of stocks that go up and down each week. Businesses that sell personal computers are interested in the proportion of households in the United States that own personal computers. Confidence intervals can be calculated for the true proportion of stocks that go up or down each week and for the true proportion of households in the United States that own personal computers.A confidence interval for a population proportion is based on the fact that the sample proportions follow an approximately normal distribution when both [latex]n \times p \geq 5[/latex] and [latex]n \times (1-p) \geq 5[/latex]. Similar to confidence intervals for population means, a confidence interval for a population proportion is constructed by taking a sample of size [latex]n[/latex] from the population, calculating the sample proportion [latex]\hat{p}[/latex], and then adding and subtracting the margin of error from [latex]\hat{p}[/latex] to get the limits of the confidence interval.In order to construct a confidence interval for a population proportion, we must be able to assume the sample proportions follow a normal distribution. As we have seen previously, we can assume the sample proportions follow a normal distribution when both [latex]n \times p \geq 5[/latex] and [latex]n \times (1-p) \geq 5[/latex]. But in this situation, the population proportion [latex]p[/latex] is unknown so we cannot check the values of [latex]n \times p[/latex] and [latex]n \times (1-p)[/latex]. Because we must take a sample and calculate the sample proportion [latex]\hat{p}[/latex], we can check the quantities [latex]n \times \hat{p}[/latex] and [latex]n \times (1-\hat{p})[/latex]. For the confidence interval, if both [latex]n \times \hat{p} \geq 5[/latex] and [latex]n \times (1-\hat{p}) \geq 5[/latex], we can assume the sample proportions follow a normal distribution.Calculating the Margin of ErrorThe margin of error for a confidence interval with confidence level [latex]C[/latex] for an unknown population proportion [latex]p[/latex] is[latex]\displaystyle{\mbox{Margin of Error}=z \times \sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}}}[/latex]where [latex]z[/latex] is the the [latex]z[/latex]-score so the area the left of [latex]z[/latex] is [latex]\displaystyle{C+\frac{1-C}{2}}[/latex]. In the margin of error formula, the sample proportion [latex]\hat{p}[/latex] is used to estimate the unknown population proportion [latex]p[/latex]. The estimated sample proportion [latex]\hat{p}[/latex] is used because [latex]p[/latex] is the unknown quantity we are trying to estimate with the confidence interval. The sample proportion [latex]\hat{p}[/latex] is calculated from the sample taken to construct the confidence interval where[latex]\displaystyle{\hat{p}=\frac{\mbox{number of items in the sample with characteristic of interest}}{n}}[/latex]Constructing the Confidence IntervalThe limits for the confidence interval with confidence level [latex]C[/latex] for an unknown population proportion [latex]p[/latex] are[latex]\begin{eqnarray*} \\ \mbox{Lower Limit} & = & \hat{p}-z \times \sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}} \\ \\ \mbox{Upper Limit} & = & \hat{p}+z \times \sqrt{\frac{\hat{p} \times (1-\hat{p})}{n}} \\ \\ \end{eqnarray*}[/latex]where [latex]z[/latex] is the [latex]z[/latex]-score so the area to the left of [latex]z[/latex] is [latex]\displaystyle{C+\frac{1-C}{2}}[/latex]. The confidence interval can only be used if we can assume the sample proportions follow a normal distribution. This means we must check that [latex]n \times \hat{p} \geq 5[/latex] and [latex]n \times (1-\hat{p}) \geq 5[/latex] before constructing the confidence interval. If one of [latex]n \times \hat{p}[/latex] or [latex]n \times (1-\hat{p})[/latex] is less than 5, we cannot construct the confidence interval.To find the [latex]z[/latex]-score to construct a confidence interval with confidence level [latex]C[/latex], use the norm.s.inv(area to the left of z) function.For area to the left of z, enter the entire area to the left of the [latex]z[/latex]-score you are trying to find. For a confidence interval, the area to the left of [latex]z[/latex] is [latex]\displaystyle{C+\frac{1-C}{2}}[/latex].The output from the norm.s.inv function is the value of [latex]z[/latex]-score needed to construct the confidence interval. The norm.s.inv function requires that we enter the entire area to the left of the unknown [latex]z[/latex]-score. This area includes the confidence level (the area in the middle of the distribution) plus the remaining area in the left tail.Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes they own cell phones.Construct a 95% confidence interval for the proportion of adult residents of this city who have cell phones.Interpret the confidence interval found in part 1.Is it reasonable to conclude that 85% of the adult residents of this city have cell phones? Explain.Solution: The sample proportion is [latex]\displaystyle{\hat{p}=\frac{421}{500}=0.842}[/latex]. We need to check [latex]n \times \hat{p}[/latex] and [latex]n \times (1-\hat{p})[/latex]:[latex]\begin{eqnarray*} \\ n \times \hat{p} & = & 500 \times 0.842=421 \geq 5 \\ \\ n \times (1-\hat{p}) & = & 500 \times (1-0.842)=79\geq 5 \\ \\ \end{eqnarray*}[/latex]Because both [latex]n \times \hat{p} \geq 5[/latex] and [latex]n \times (1-\hat{p}) \geq 5[/latex], the sample proportions follow a normal distribution and we can construct the confidence interval.To find the confidence interval, we need to find the [latex]z[/latex]-score for the 95% confidence interval. This means that we need to find the [latex]z[/latex]-score so that the entire area to the left of [latex]z[/latex] is [latex]\displaystyle{0.95+\frac{1-0.95}{2}=0.975}[/latex].Functionnorm.s.invAnswerField 10.9751.9599So [latex]z=1.9599....[/latex]. The 95% confidence interval is[latex]\begin{eqnarray*}\\\mbox{Lower Limit}&=&\hat{p}-z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.842-1.9599...\times\sqrt{\frac{0.842\times(1-0.842)}{500}}\\&=&0.8100\\\\ \mbox{Upper Limit}&=&\hat{p}+z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.842+1.9599...\times\sqrt{\frac{0.842\times(1-0.842)}{500}}\\&=&0.8740\\\\\end{eqnarray*}[/latex]We are 95% confident that the proportion of adult residents of this city who have cell phones is between 81% and 87.4%.It is reasonable to conclude that 85% of the adult residents of this city have cell phones because 85% is inside the confidence interval.When calculating the limits for the confidence interval keep all of the decimals in the [latex]z[/latex]-score and other values throughout the calculation. This will ensure that there is no round-off error in the answers. You can use Excel to do the calculation of the limits, clicking on the cells containing the [latex]z[/latex]-score and any other values, to ensure that all of the decimal places are used in the calculation.The limits for the confidence interval are percents. For example, the upper limit of 0.8740 is the decimal form of a percent: 87.4%.When writing down the interpretation of the confidence interval, make sure to include the confidence level, the actual population proportion captured by the confidence interval (i.e. be specific to the context of the question), and express the limits as percents.95% of all confidence interval constructed this way contain the proportion of adult residents in this city that have a cell phone. For example, if we constructed 100 of these confidence (using 100 different samples of size 500), we would expect 95 of them to contain the true proportion of adult residents in this city that have a cell phone.Suppose 250 randomly selected people are surveyed to determine if they own a tablet. Of the 250 surveyed, 98 reported owning a tablet.Construct a 94% confidence interval for the proportion of people who own tablets.Interpret the confidence interval found in part 1.Is it reasonable to assume that 30% of people own tablets? Explain.Click to see SolutionFunctionnorm.s.invAnswerField 10.971.8807[latex]\begin{eqnarray*}\\\mbox{Lower Limit}&=&\hat{p}-z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.392-1.8807...\times\sqrt{\frac{0.392\times(1-0.392)}{250}}\\&=&0.3339\\\\\mbox{Upper Limit}&=&\hat{p}+z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.392+1.8807...\times\sqrt{\frac{0.392\times(1-0.392)}{250}}\\&=&0.4501\\\\\end{eqnarray*}[/latex]We are 94% confident that the proportion of people who own tablets is between 33.39% and 45.01%.It is not reasonable to claim the proportion of people who own tablets is 30% because 30% is outside the confidence interval.For a class project, a political science student at a large university wants to estimate the percent of students who are registered voters. He surveys 500 students and finds that 300 are registered voters.Construct a 90% confidence interval for the percent of students who are registered voters.Interpret the confidence interval found in part 1.Solution:The sample proportion is [latex]\displaystyle{\hat{p}=\frac{300}{500}=0.6}[/latex]. We need to check [latex]n \times \hat{p}[/latex] and [latex]n \times (1-\hat{p})[/latex]:[latex]\begin{eqnarray*}\\\times\hat{p}&=&500\times0.6=300\geq 5\\\\ n\times(1-\hat{p})&=&500\times(1-0.6)=200\geq 5\\\\\end{eqnarray*}[/latex]Because both [latex]n \times \hat{p} \geq 5[/latex] and [latex]n \times (1-\hat{p}) \geq 5[/latex], the sample proportions follow a normal distribution and we can construct the confidence interval.To find the confidence interval, we need to find the [latex]z[/latex]-score for the 90% confidence interval. This means that we need to find the [latex]z[/latex]-score so that the entire area to the left of [latex]z[/latex] is [latex]\displaystyle{0.90+\frac{1-0.90}{2}=0.95}[/latex].Functionnorm.s.invAnswerField 10.951.6448So [latex]z=1.6448....[/latex]. The 90% confidence interval is[latex]\begin{eqnarray*}\\\mbox{Lower Limit}&=&\hat{p}-z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.6-1.6448...\times \sqrt{\frac{0.6\times(1-0.6)}{500}}\\&=&0.5640\\\\\mbox{Upper Limit}&=&\hat{p}+z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.6+1.6448...\times\sqrt{\frac{0.6\times(1-0.6)}{500}}\\&=&0.6360\\\\\end{eqnarray*}[/latex]We are 90% confident that the percent of students who are registered voters is between 56.4% and 63.6%.A student polls her school to see if students in the school district are for or against the new legislation regarding school uniforms. She surveys 600 students and finds that 480 are against the new legislation.Construct a 98% confidence interval for the proportion of students who are against the new legislation.Interpret the confidence interval found in part 1.A parents group claims that only 75% of students are against the legislation. Is it reasonable for the group to make this claim? Explain.Click to see SolutionFunctionnorm.s.invAnswerField 10.992.3263[latex]\begin{eqnarray*}\\\mbox{Lower Limit}&=&\hat{p}-z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.8-2.3264...\times\sqrt{\frac{0.8 \times(1-0.8)}{600}}\\&=&0.7620\\\\\mbox{Upper Limit}&=&\hat{p}+z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\&=&0.8 +2.3263...\times\sqrt{\frac{0.8\times(1-0.8)}{600}}\\&=&0.8380\\\\\end{eqnarray*}[/latex]We are 98% confident that the proportion of students who are against the new legislation is between 76.20% and 83.80%.It is not reasonable for the group to claim the proportion is 75% because 75% is outside of the confidence interval.Watch this video: Confidence Interval for a population proportion by Excel is Fun [8:34] Watch this video: Confidence Interval for a population proportion by Excel is Fun [4:51] Concept ReviewSome statistical measures, like many survey questions, measure qualitative rather than quantitative data. In this case, the population parameter being estimated is a proportion. It is possible to create a confidence interval for the true population proportion following procedures similar to those used in creating confidence intervals for population means. The formulas are slightly different, but they follow the same reasoning.The general form for a confidence interval for a single population proportion is given by[latex]\begin{eqnarray*}\\\mbox{Lower Limit}&=&\hat{p}-\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\\\\mbox{Upper Limit}&=&\hat{p}+z\times\sqrt{\frac{\hat{p}\times(1-\hat{p})}{n}}\\\\\end{eqnarray*}[/latex]where [latex]z[/latex] is the the [latex]z[/latex]-score so the area to the left of [latex]z[/latex] is [latex]\displaystyle{C+\frac{1-C}{2}}[/latex]. Attribution8.3A Population Proportion in Introductory Statistics by OpenStaxis licensed under aCreative Commons Attribution 4.0 International License. By the end of this chapter, the student should be able to: Calculate and interpret confidence intervals for estimating a population proportion. Calculate the sample size required to estimate a population proportion given a desired confidence level and margin of error. During an election year, we see articles in the newspaper that state confidence intervals in terms of proportions or percentages. For example, a poll for a particular candidate running for president might show that the candidate has 40% of the vote within three percentage points (if the sample is large enough). Often, election polls are calculated with 95% confidence, so, the pollsters would be 95% confident that the true proportion of voters who favored the candidate would be between 0.37 and 0.43: (0.40 0.03,0.40 + 0.03). Investors in the stock market are interested in the true proportion of stocks that go up and down each week. Businesses that sell personal computers are interested in the proportion of households in the United States that own personal computers. Confidence intervals can be calculated for the true proportion of stocks that go up or down each week and for the true proportion of households in the United States that own personal computers. The procedure to find the confidence interval, the sample size, the error bound, and the confidence level for a proportion is similar to that for the population mean, but the formulas are different. How do you know you are dealing with a proportion problem? There is no mention of a mean or average! To form a proportion, take \(X\), the random variable for the number of successes and divide it by \(n\), the number of trials (or the sample size). The random variable \(\hat{P}\) (read "P hat") is that proportion, \[\hat{P}= \dfrac{X}{n}onumber \] When \(n\) is large and \(p\) is not close to zero or one, we can use the normal distribution to approximate the number of successes. \[X \sim N(np, \sqrt{npq})onumber \] If we divide the random variable, the mean, and the standard deviation by \(n\), we get a normal distribution of proportions with \(P \), called the estimated proportion, as the random variable. (Recall that a proportion as the number of successes divided by \(n\).) \[\dfrac{X}{n} = \hat{P}\sim N\left(\dfrac{np}{n}, \dfrac{\sqrt{npq}}{n}\right)onumber \] Using algebra to simplify: \[\dfrac{\sqrt{npq}}{n} = \sqrt{\dfrac{pq}{n}}onumber \] \(\hat{P}\) follows a normal distribution for proportions: \[\dfrac{X}{n} = \hat{P}\sim N\left(\dfrac{np}{n}, \dfrac{\sqrt{npq}}{n}\right)onumber \] The error bound (EBP) for a proportion is \[EBP = \left(z_{\frac{\alpha}{2}}\right)\left(\sqrt{\dfrac{\hat{p}\hat{q}}{n}}\right)onumber \] where \(\hat{q}\ = 1 - \hat{p}\). This formula is similar to the error bound formula for a mean, except that the "appropriate standard deviation" is different. For a mean, when the population standard deviation is known, the appropriate standard deviation that we use is \(\dfrac{\sigma}{\sqrt{n}}\). For a proportion, the appropriate standard deviation is \[\sqrt{\dfrac{pq}{n}}.onumber \] However, in the error bound formula, we use \[\sqrt{\dfrac{\hat{p}\hat{q}}{n}}onumber \] as the standard deviation, instead of \[\sqrt{\dfrac{pq}{n}}.onumber \] In the error bound formula, the sample proportions \(\hat{p}\) and\(\hat{q}\) are estimates of the unknown population proportions p and q. The estimated proportions \(\hat{p}\) and \(\hat{q}\) are used because \(p\) and \(q\) are not known. The sample proportions \(\hat{p}\) and \(\hat{q}\) are calculated from the data: \(\hat{p}\) is the estimated proportion of successes, and \(\hat{q}\) is the estimated proportion of failures. There are many online calculators that can be used to compute the Margin of Error. For example, you can use this one: Also,you are encouraged to ask your instructor about which calculator is allowed/recommended for this course. Use the calculatorprovided above to verify the following statements: When \(\alpha=0.1, n=200, \hat{p}=0.43\) the EBPis \(0.0577\) When \(\alpha=0.05, n=100, \hat{p}=0.81\) the EBPis \(0.0768\) When \(\alpha=0.01, n=250, \hat{p}=0.57\) the EBPis \(0.0806\) Find EBPwhen \(\alpha=0.07, n=168, \hat{p}=0.82\). Answer 0.0806 The confidence interval has the form \[(\hat{p} EBP, \hat{p}+ EBP).onumber \] where \(EBP\) is error bound for the proportion. \(\hat{p}= \dfrac{x}{n}\) \(\hat{p}=\) the estimated proportion of successes (\hat{p}is a point estimate for p, the true proportion.) \(x =\) the number of successes \(n =\) the size of the sample The confidence interval can be used only if the number of successes \(n\hat{p}\) and the number of failures \(n\hat{q}\) are both greater than five. The graph gives a picture of the entire situation. \[CL + \dfrac{\alpha}{2} + \dfrac{\alpha}{2} = CL + \alpha = 1.onumber \] Figure 8.2.2. There are many online calculators that can be used to compute the confidence intervals. For example, you can use this one: Also,you are encouraged to ask your instructor about which calculator is allowed/recommended for this course. Use the calculatorprovided above to verify the following: Confidence Level (%): \(95\) Sample Size: \(197\) Number of Successes: \(61\) 95% Confidence Interval: \((0.2450,0.3742)\) Find a90% confidence interval when the sample size is 250 and thenumber of successes is85. Answer (0.2907, 0.3893) The interpretation should clearly state the confidence level (\(CL\)), explain what population parameter is being estimated (here, a population proportion), and state the confidence interval (both endpoints). "We estimate with ___% confidence that the true population proportoin(include the context of the problem) is between ___ and ___ ." Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes - they own cell phones. Using a 95% confidence level, compute a confidence interval estimate for the true proportion of adult residents of this city who have cell phones. Solution To calculate the confidence interval, you must find \(\hat{p}\), \(\hat{q}\), and \(EBP\). \(n = 500\) \(x =\) the number of successes \(= 421\) \[\hat{p} = \dfrac{x}{n} = \dfrac{421}{500} = 0.842onumber \] \(\hat{p} = 0.842\) is the sample proportion; this is the point estimate of the population proportion. \[\hat{q} = 1 \hat{p} = 1 0.842 = 0.158onumber \] Since \(CL = 0.95\), then \[\alpha = 1 CL = 1 0.95 = 0.05\left(\dfrac{\alpha}{2}\right) = 0.025.onumber \] Then \[z_{\dfrac{\alpha}{2}} = z_{0.025 = 1.96}onumber \] Use the TI-83, 83+, or 84+ calculator command invNorm(0.975,0,1) to find \(z_{0.025}\). Remember that the area to the right of \(z_{0.025}\) is \(0.025\) and the area to the left of \(z_{0.025}\) is \(0.975\). This can also be found using appropriate commands on other calculators, using a computer, or using a Standard Normal probability table. \[EBP = \left(z_{\dfrac{\alpha}{2}}\right)\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = (1.96)\sqrt{\dfrac{(0.842)(0.158)}{500}} = 0.032onumber \] \[\hat{p} EBP = 0.842 0.032 = 0.81onumber \] \[\hat{p} + EBP = 0.842 + 0.032 = 0.874onumber \] The confidence interval for the true binomial population proportion is \((\hat{p} EBP, \hat{p} +EBP) = (0.810, 0.874)\). Interpretation:We estimate with 95% confidence that between 81% and 87.4% of all adult residents of this city have cell phones. Explanation of 95% Confidence Level:Ninety-five percent of the confidence intervals constructed in this way would contain the true value for the population proportion of all adult residents of this city who have cell phones. Suppose 250 randomly selected people are surveyed to determine if they own a tablet. Of the 250 surveyed, 98 reported owning a tablet. Using a 95% confidence level, compute a confidence interval estimate for the true proportion of people who own tablets. Answer (0.3315, 0.4525) For a class project, a political science student at a large university wants to estimate the percent of students who are registered voters. He surveys 500 students and finds that 300 are registered voters. Compute a 90% confidence interval for the true percent of students who are registered voters, and interpret the confidence interval. Answer \(x = 300\) and\(n = 500\) \[\hat{p} = \dfrac{x}{n} = \dfrac{300}{500} = 0.600onumber \] \[\hat{q} = 1 \hat{p} = 1 0.600 = 0.400onumber \] Since \(CL = 0.90\), then \[\alpha = 1 CL = 1 0.90 = 0.10\left(\dfrac{\alpha}{2}\right) = 0.05\] \[z_{\dfrac{\alpha}{2}} = z_{0.05} = 1.645onumber \] Use the TI-83, 83+, or 84+ calculator command invNorm(0.95,0,1) to find \(z_{0.05}\). Remember that the area to the right of \(z_{0.05}\) is 0.05 and the area to the left of \(z_{0.05}\) is 0.95. This can also be found using appropriate commands on other calculators, using a computer, or using a standard normal probability table. \[EBP = \left(z_{\dfrac{\alpha}{2}}\right)\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = (1.645)\sqrt{\dfrac{(0.60)(0.40)}{500}} = 0.036onumber \] \[\hat{p} EBP = 0.60 0.036 = 0.564onumber \] \[\hat{p} + EBP = 0.60 + 0.036 = 0.636onumber \] The confidence interval for the true binomial population proportion is \((\hat{p} EBP, \hat{p} +EBP) = (0.564,0.636)\). Interpretation We estimate with 90% confidence that the true percent of all students that are registered voters is between 56.4% and 63.6%. Alternate Wording: We estimate with 90% confidence that between 56.4% and 63.6% of ALL students are registered voters. Explanation of 90% Confidence Level:Ninety percent of all confidence intervals constructed in this way contain the true value for the population percent of students that are registered voters. A student polls his school to see if students in the school district are for or against the new legislation regarding school uniforms. She surveys 600 students and finds that 480 are against the new legislation. Compute a 90% confidence interval for the true percent of students who are against the new legislation, and interpret the confidence interval. In a sample of 300 students, 68% said they own an iPod and a smart phone. Compute a 97% confidence interval for the true percent of students who own an iPod and a smartphone. Answer (0.7731, 0.8269); We estimate with 90% confidence that the true percent of all students in the district who are against the new legislation is between 77.31% and 82.69% A random sample of 29statistics students was asked: Have you smoked a cigarette in the past week? Eight students reported smoking within the past week.Find a 95% confidence interval for the true proportion of statistics students who smoke. Solution Eightstudents out of 29reported smoking within the past week, so \(x = 8\) and \(n = 29\). \[\hat{p}= \dfrac{x}{n} = \dfrac{8}{29} \approx 0.276onumber \] \[\hat{q} = 1 \hat{p} = 1 0.276 = 0.724onumber \] Since \(CL = 0.95\), we know \(\alpha = 1 0.95 = 0.05\) and \(\dfrac{\alpha}{2} = 0.025\). \[z_{0.025} = 1.96onumber \] \(EPB = \left(z_{\dfrac{\alpha}{2}}\right)\sqrt{\dfrac{\hat{p}\hat{q}}{n}} = (1.96)\sqrt{\dfrac{0.276(0.724)}{29}} \approx 0.163\) \[\hat{p} EPB = 0.276 0.163 = 0.113onumber \] \[\hat{p} + EPB = 0.276 + 0.163 = 0.439onumber \] We are 95% confident that the true proportion of all statistics students who smoke cigarettes is between 0.113 and 0.439. Out of a random sample of 69freshmen at State University, 33students have declared a major. Find a 96% confidence interval for the true proportion of freshmen at State University who have declared a major. Solution We have \(x = 33\) and \(n = 69\). \[\hat{p}= 33/69 \approx 0.478onumber \] \[\hat{q} = 1 \hat{p} = 1 0.478 = 0.522onumber \] Since \(CL = 0.96\), we know \(\alpha = 1 0.96 = 0.04\) and \(\dfrac{\alpha}{2} = 0.02\). \[z_{0.02} = 2.054onumber \] \[EPB = \left(z_{\dfrac{\alpha}{2}}\right)\sqrt{\dfrac{p'q'}{n}} = (2.054)\left(\sqrt{\dfrac{(0.478)(0.522)}{69}}\right) - 0.124onumber \] \[\hat{p} EPB = 0.478 0.124 = 0.354onumber \] \[\hat{p} + EPB = 0.478 + 0.124 = 0.602onumber \] We are 96% confident that between 35.4% and 60.2% of all freshmen at State U have declared a major. The Berkman Center for Internet & Society at Harvard recently conducted a study analyzing the privacy management habits of teen internet users. In a group of 50 teens, 13 reported having more than 500 friends on Facebook. Use the plus four method to find a 90% confidence interval for the true proportion of teens who would report having more than 500 Facebook friends. Solution A Using plus-four, we have \(x = 13 + 2 = 15\) and \(n = 50 + 4 = 54\). \[p = 1554 \approx 0.278onumber \] \[q = 1 p = 1 0.241 = 0.722onumber \] Since \(CL = 0.90\), we know \(\alpha = 1 0.90 = 0.10\) and \(\dfrac{\alpha}{2} = 0.05\). \[z_{0.05} = 1.645onumber \] \[EPB = \left(z_{\dfrac{\alpha}{2}}\right)\left(\sqrt{\dfrac{p'q'}{n}}\right) = (1.645)\left(\sqrt{\dfrac{(0.278)(0.722)}{54}}\right) \approx 0.100onumber \] \[p EPB = 0.278 0.100 = 0.178onumber \] \[p + EPB = 0.278 + 0.100 = 0.378onumber \] We are 90% confident that between 17.8% and 37.8% of all teens would report having more than 500 friends on Facebook. Solution B Press STAT and arrow over to TESTS. Arrow down to A:1-PropZint. Press ENTER.Arrow down to \(x\) and enter 15.Arrow down to \(n\) and enter 54.Arrow down to C-Level and enter 0.90.Arrow down to Calculate and press ENTER. The confidence interval is (0.178, 0.378). If researchers desire a specific margin of error, then they can use the error bound formula to calculate the required sample size. The error bound formula for a population proportion is \[EBP = \left(z_{\frac{\alpha}{2}}\right)\left(\sqrt{\dfrac{\hat{p}\hat{q}}{n}}\right)onumber \] Solving for \(n\) gives you an equation for the sample size. \[n = \dfrac{\left(z_{\frac{\alpha}{2}}\right)^{2}(\hat{p}\hat{q})}{EBP^{2}}onumber \] Suppose a mobile phone company wants to determine the current percentage of customers aged 50+ who use text messaging on their cell phones. How many customers aged 50+ should the company survey in order to be 90% confident that the estimated (sample) proportion is within three percentage points of the true population proportion of customers aged 50+ who use text messaging on their cell phones. Answer From the problem, we know that \(\bf{EBP = 0.03}\) (3%=0.03) and \(z_{\dfrac{\alpha}{2}} z_{0.05} = 1.645\) because the confidence level is 90%. However, in order to find \(n\), we need to know the estimated (sample) proportion \(\hat{p}\). Remember that \(\hat{q} = 1 \hat{p}\). But, we do not know \(\hat{p}\) yet. Since we multiply \(\hat{p}\) and \(\hat{q}\) together, we make them both equal to 0.5 because \(\hat{p}\hat{q} = (0.5)(0.5) = 0.25\) results in the largest possible product. (Try other products: \((0.6)(0.4) = 0.24\); \((0.3)(0.7) = 0.21\); \((0.2)(0.8) = 0.16\) and so on). The largest possible product gives us the largest \(n\). This gives us a large enough sample so that we can be 90% confident that we are within three percentage points of the true population proportion. To calculate the sample size \(n\), use the formula and make the substitutions. \[n = \dfrac{z^{2}\hat{p}\hat{q}}{EBP^{2}}onumber \] gives \[n = \dfrac{1.645^{2}(0.5)(0.5)}{0.03^{2}} = 751.7onumber \] Round the answer to the next higher value. The sample size should be 752 cell phone customers aged 50+ in order to be 90% confident that the estimated (sample) proportion is within three percentage points of the true population proportion of all customers aged 50+ who use text messaging on their cell phones. Suppose an internet marketing company wants to determine the current percentage of customers who click on ads on their smartphones. How many customers should the company survey in order to be 90% confident that the estimated proportion is within five percentage points of the true population proportion of customers who click on ads on their smartphones? Answer 271 customers should be surveyed. Check the Real Estate section in your local A confidence interval estimates a population proportion using a range based on a sample proportion.The sample proportion, called p-hat, is used to estimate the unknown population proportion, p.The margin of error in a confidence interval uses a z-score and the standard error of p-hat. Confidence intervals can be used to estimate several population parameters. One type of parameter that can be estimated using inferential statistics is a population proportion. For example, we may want to know the percentage of the U.S. population who supports a particular piece of legislation. For this type of question, we need to find a confidence interval. In this article, we will see how to construct a confidence interval for a population proportion, and examine some of the theory behind this. We begin by looking at the big picture before we get into the specifics. The type of confidence interval that we will consider is of the following form: Estimate +/- Margin of Error This means that there are two numbers that we will need to determine. These values are an estimate for the desired parameter, along with the margin of error. Before conducting any statistical test or procedure, it is important to make sure that all of the conditions are met. For a confidence interval for a population proportion, we need to make sure that the following hold:We have a simple random sample of size n from a large populationOur individuals have been chosen independently of one another.There are at least 15 successes and 15 failures in our sample. If the last item is not satisfied, then it may be possible to adjust our sample slightly and to use a plus-four confidence interval. In what follows, we will assume that all of the above conditions have been met. We start with the estimate for our population proportion. Just as we use a sample mean to estimate a population mean, we use a sample proportion to estimate a population proportion. The population proportion is an unknown parameter. The sample proportion is a statistic. This statistic is found by counting the number of successes in our sample and then dividing by the total number of individuals in the sample. The population proportion is denoted by p and is self-explanatory. The notation for the sample proportion is a little more involved. We denote a sample proportion as p, and we read this symbol as "p-hat" because it looks like the letter p with a hat on top. This becomes the first part of our confidence interval. The estimate of p is p. To determine the formula for the margin of error, we need to think about the sampling distribution of p. We will need to know the mean, the standard deviation, and the particular distribution that we are working with. The sampling distribution ofp is a binomial distribution with probability of success p and n trials. This type of random variable has a mean of p and standard deviation of (p(1 - p)/n)0.5. There are two problems with this. The first problem is that a binomial distribution can be very tricky to work with. The presence of factorials can lead to some very large numbers. This is where the conditions help us. As long as our conditions are met, we can estimate the binomial distribution with the standard normal distribution. The second problem is that the standard deviation ofp uses p in its definition. The unknown population parameter is to be estimated by using that very same parameter as a margin of error. This circular reasoning is a problem that needs to be fixed. The way out of this conundrum is to replace the standard deviation with its standard error. Standard errors are based upon statistics, not parameters. A standard error is used to estimate a standard deviation.What makes this strategy worthwhile is that we no longer need to know the value of the parameter p. To use the standard error, we replace the unknown parameter p with the statistic p. The result is the following formula for a confidence interval for a population proportion: p +/- z* (p(1 - p)/n)0.5. Here the value of z* is determined by our level of confidence C.For the standard normal distribution, exactly C percent of the standard normal distribution is between -z* and z*.Common values for z* include 1.645 for 90% confidence and 1.96 for 95% confidence. Let's see how this method works with an example.Suppose that we wish to know with 95% confidence the percent of the electorate in a county that identifies itself as Democratic. We conduct a simple random sample of 100 people in thiscounty and find that 64 of them identify as a Democrat. We see that all of the conditions are met.The estimate of our population proportion is 64/100 = 0.64.This is the value of the sample proportion p, and it is the center of our confidence interval. The margin of error is comprised of two pieces.The first is z*.As we said, for 95% confidence, the value of z* = 1.96. The other part of the margin of error is given by the formula (p(1 - p)/n)0.5.We set p = 0.64 and calculate = the standard error to be (0.64(0.36)/100)0.5 = 0.048. We multiply these two numbers together and obtain a margin of error of 0.09408.The end result is: 0.64 +/- 0.09408, or we can rewrite this as 54.592% to 73.408%.Thus we are 95% confident that the true population proportion of Democrats is somewhere in the range of these percentages.This means that in the long run, our technique and formula will capture the population proportion of 95% of the time. There are a number of ideas and topics that are connected to this type of confidence interval.For instance, we could conduct a hypothesis test pertaining to the value of the population proportion.We could also compare two proportions from two different populations. During an election year, we see articles in the newspaper that state confidence intervals in terms of proportions or percentages. For example, a poll for a particular candidate running for president might show that the candidate has 40% of the vote within three percentage points (if the sample is large enough). Often, election polls are calculated with 95% confidence, so, the pollsters would be 95% confident that the true proportion of voters who favored the candidate would be between 0.37 and 0.43: (0.40 0.03,0.40 + 0.03). Investors in the stock market are interested in the true proportion of stocks that go up and down each week. Businesses that sell personal computers are interested in the proportion of households in the United States that own personal computers. Confidence intervals can be calculated for the true proportion of stocks that go up or down each week and for the true proportion of households in the United States that own personal computers. The build a confidence interval for population proportion \(p\), we use: \[ {\hat p} z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{ {\hat p}(1-{\hat p})}{n}} < p< {\hat p} + z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{ {\hat p}(1-{\hat p})}{n}}. \] where \( {\hat p} = \dfrac{x}{n}\), the estimated proportion of successes \( {\hat p}\) is a point estimate for \(p\), the true proportion.) \(x =\) the number of successes \(n =\) the size of the sample \(z_{\frac{\alpha}{2}}\cdot\sqrt{\dfrac{ {\hat p}(1-{\hat p})}{n}} \) is called the margin of error In the margin of error formula, the sample proportions \( {\hat p}\) and 1-{\hat p} are estimates of the unknown population proportions p and 1-p. The estimated proportions \({\hat p}\) and \(1-{\hat p}\) are used because \(p\) and \(1-{\hat p}\) are not known. The sample proportions \({\hat p}\) and \(1-{\hat p}\) are calculated from the data: \({\hat p}\) is the estimated proportion of successes, and \(1-{\hat p}\) is the estimated proportion of failures. Example 7.2.1 Suppose that a market research firm is hired to estimate the percent of adults living in a large city who have cell phones. Five hundred randomly selected adult residents in this city are surveyed to determine whether they have cell phones. Of the 500 people surveyed, 421 responded yes - they own cell phones. Using a 95% confidence level, compute a confidence interval estimate for the true proportion of adult residents of this city who have cell phones. Answer The first solution is step-by-step (Solution A). The second solution uses a function of the TI-83, 83+, or 84 calculators (Solution B). Solution A \(n = 500\) \(x =\) the number of successes \(= 421\) \[{\hat p} = \dfrac{x}{n} = \dfrac{421}{500} = 0.842onumber \] \({\hat p} = 0.842\) is the sample proportion; this is the point estimate of the population proportion. \[1 {\hat p} = 1 0.842 = 0.158onumber \] Since the confidence level \(CL = 0.95\), then \(\alpha = 1 CL = 1 0.95 = 0.05 \) So, \( \dfrac{\alpha}{2} = 0.025.onumber \) Then \[z_{\frac{\alpha}{2}} = z_{0.025} = 1.96 onumber \] Use the TI-83, 83+, or 84+ calculator command invNorm(0.975,0,1) to find \(z_{0.025}\). Remember that the area to the right of \(z_{0.025}\) is \(0.025\) and the area to the left of \(z_{0.025}\) is \(0.975\). This can also be found using appropriate commands on other calculators, using a computer, or using a Standard Normal probability table. margin of error \(=z_{\frac{\alpha}{2}} \cdot \sqrt{ \dfrac{{\hat p}(1-{\hat p})}{n} } \) \(= 1.96 \cdot\sqrt{\dfrac{(0.842)(0.158)}{500}} = 0.032 \) \[{\hat p} \text{margin of error }= 0.842 0.032 = 0.81onumber \] \[{\hat p} + \text{ margin of error } = 0.842 + 0.032 = 0.874onumber \] The confidence interval for the true binomial population proportion is \(({\hat p} \text{margin of error}, {\hat p}+\text{margin of error}) = (0.810, 0.874)\). Interpretation We estimate with 95% confidence that between 81% and 87.4% of all adult residents of this city have cell phones. Explanation of 95% Confidence Level Ninety-five percent of the confidence intervals constructed in this way would contain the true value for the population proportion of all adult residents of this city who have cell phones. Solution B Press STAT and arrow over to TESTS. Arrow down to A:1-PropZint. Press ENTER.Arrow down to xx and enter 421.Arrow down to nn and enter 500.Arrow down to C-Level and enter .95.Arrow down to Calculate and press ENTER.The confidence interval is (0.81003, 0.87397). Example 7.2.2 For a class project, a political science student at a large university wants to estimate the percent of students who are registered voters. He surveys 500 students and finds that 300 are registered voters. Compute a 90% confidence interval for the true percent of students who are registered voters, and interpret the confidence interval. Answer The first solution is step-by-step (Solution A). The second solution uses a function of the TI-83, 83+, or 84 calculators (Solution B). Solution A \(x = 300\) and \(n = 500\) \[{\hat p} = \dfrac{x}{n} = \dfrac{300}{500} = 0.600onumber \] \[1-{\hat p} = 1 0.600 = 0.400onumber \] Since \(CL = 0.90\), then \(\alpha = 1 CL = 1 0.90 = 0.10 \) So, \( \dfrac{\alpha}{2} = 0.05.onumber \) \[z_{\dfrac{\alpha}{2}} = z_{0.05} = 1.645onumber \] Use the TI-83, 83+, or 84+ calculator command invNorm(0.95,0,1) to find \(z_{0.05}\). Remember that the area to the right of \(z_{0.05}\) is 0.05 and the area to the left of \(z_{0.05}\) is 0.95. This can also be found using appropriate commands on other calculators, using a computer, or using a standard normal probability table. margin of error \(=z_{\frac{\alpha}{2}} \cdot \sqrt{ \dfrac{{\hat p}(1-{\hat p})}{n} } \) \(= 1.645 \cdot \sqrt{\dfrac{(0.60)(0.40)}{500}} = 0.036 \) \[{\hat p}-\text{ margin of error } = 0.60 0.036 = 0.564onumber \] \[{\hat p} + \text{margin of error} = 0.60 + 0.036 = 0.636onumber \] The confidence interval for the true binomial population proportion is \(({\hat p} \text{margin of error}, {\hat p}+\text{margin of error}) = (0.564,0.636)\). Interpretation We estimate with 90% confidence that the true percent of all students that are registered voters is between 56.4% and 63.6%. Alternate Wording: We estimate with 90% confidence that between 56.4% and 63.6% of ALL students are registered voters. Explanation of 90% Confidence Level Ninety percent of all confidence intervals constructed in this way contain the true value for the population percent of students that are registered voters. Solution B Press STAT and arrow over to TESTS. Arrow down to A:1-PropZint. Press ENTER.Arrow down to xx and enter 300.Arrow down to nn and enter 500.Arrow down to C-Level and enter 0.90.Arrow down to Calculate and press ENTER. The confidence interval is (0.564, 0.636). Example 7.2.3 To estimate the proportion of students at a large college who are female, a random sample of \(120\) students is selected. There are \(69\) female students in the sample. Construct a \(90\%\) confidence interval for the proportion of all students at the college who are female. Solution A The proportion of students in the sample who are female is \[ \hat{p} =69/120=0.575 onumber\] Confidence level \(90\%\) means that \(\alpha =1-0.90=0.10\) so \(\alpha /2=0.05\). From the last line of Figure 7.1.6 we obtain \(z_{0.05}=1.645\). Thus \[\hat{p}\pm z_{\alpha /2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}=0.575\pm 1.645\sqrt{\frac{(0.575)(0.425)}{120}}=0.575\pm 0.074 onumber\] One may be \(90\%\) confident that the true proportion of all students at the college who are female is contained in the interval \((0.575-0.074,0.575+0.074)=(0.501,0.649)\). Solution B Press STAT and arrow over to TESTS. Arrow down to A:1-PropZint. Press ENTER. Arrow down to \(x\) and enter 69. Arrow down to \(n\) and enter 120.Arrow down to C-Level and enter 0.90.Arrow down to Calculate and press ENTER. The confidence interval is (0.501,0.649). Contributors and Attributions Barbara Illowsky and Susan Dean (De Anza College) with many other contributing authors. Content produced by OpenStax College is licensed under a Creative Commons Attribution License 4.0 license. Download for free at 18.114. Explain what a confidence interval represents and determine how changes in sample size and confidence level affect the precision of the confidence interval.Find confidence intervals for the population mean and the population proportion (when certain conditions are met), and perform sample size calculations. As we mentioned in the introduction to this module, when the variable that were interested in studying in the population iscategorical, the parameter we are trying to infer about is thepopulation proportion (p)associated with that variable. We also learned that the point estimator for the population proportion p is the sample proportion[latex]\hat{p}[/latex].To refresh your memory, here is a picture that summarizes an example we looked at.We are now moving on to interval estimation of p. In other words, we would like to develop a set of intervals that, with different levels of confidence, will capture the value of p. Weve actually done all the groundwork and discussed all the big ideas of interval estimation when we talked about interval estimation for , so well be able to go through it much faster. Lets begin.Recall that the general form of any confidence interval for an unknown parameter is:estimate [latex]\pm[/latex] margin of errorSince the unknown parameter here is the population proportion p, the point estimator (as I reminded you above) is the sample proportion[latex]\hat{p}[/latex]. The confidence interval for p, therefore, has the form:[latex]\hat{\mathcal{p}}\pm\mathcal{m}[/latex](Recall that m is the notation for the margin of error.) The margin of error (m) tells us with a certain confidence what the maximum estimation error is that we are making, or in other words, that[latex]\hat{p}[/latex]is different from p (the parameter it estimates) by no more than m units.From our previous discussion on confidence intervals, we also know that the margin of error is the product of two components:m = confidence multiplier [latex]\cdot[/latex] SD of the estimatorTo figure out what these two components are, we need to go back to a result we obtained in the Sampling Distributions module of the Probability unit about the sampling distribution of[latex]\hat{p}[/latex]. We found that under certain conditions (which well come back to later),[latex]\hat{p}[/latex] has a normal distribution with mean p, and standard deviation [latex]\sqrt{\frac{p(1-p)}{n}}[/latex]. This result makes things very simple for us, because it reveals what the two components are that the margin of error is made of:* Since, like the sampling distribution of[latex]\overline{X}[/latex], the sampling distribution of[latex]\hat{p}[/latex]is normal, the confidence multipliers that well use in the confidence interval for p will be the same z* multipliers we use for the confidence interval for when is known (usingexactlythe same reasoning and the same probability results). The multipliers well use, then, are:1.645, 2, and 2.576 at the 90%, 95% and 99% confidence levels, respectively.* The standard deviation of our estimator[latex]\hat{p}[/latex]is[latex]\sqrt{\frac{\mathcal{p}\left(1-\mathcal{p}\right)}{\mathcal{n}}}[/latex]Putting it all together, we find that the confidence interval for p should be:[latex]\hat{\mathcal{p}}\pm\mathcal{z}*\ \bullet\sqrt{\frac{\mathcal{p}\left(1-\mathcal{p}\right)}{\mathcal{n}}}[/latex]We just have to solve one practical problem and were done. Were trying to estimate theunknownpopulation proportion p, so having it appear in the confidence interval doesnt make any sense. To overcome this problem, well do the obvious thingWell replace p with its sample counterpart, [latex]\hat{p}[/latex], and work with the standard error of [latex]\hat{p}[/latex], [latex]\sqrt{\frac{\mathcal{\hat{p}}\left(1-\mathcal{\hat{p}}\right)}{\mathcal{n}}}[/latex]Now were done. The confidence interval for the population proportion p is:[latex]\hat{\mathcal{p}}\pm\mathcal{z}*\ \bullet\sqrt{\frac{\mathcal{\hat{p}}\left(1-\mathcal{\hat{p}}\right)}{\mathcal{n}}}[/latex]As youll see from the examples well present in this unit, in estimating the population proportion comes up a lot in the context of polls. The drug Viagra became available in the U.S. in May, 1998, in the wake of an advertising campaign that was unprecedented in scope and intensity. A Gallup poll found that by the end of the first week in May, 643 out of a random sample of 1,005 adults were aware that Viagra was an impotency medication (based on Viagra A Popular Hit, a Gallup poll analysis by Lydia Saad, May 1998).Lets estimate the proportion p of all adults in the U.S. who by the end of the first week of May 1998 were already aware of Viagra and its purpose by setting up a 95% confidence interval for p.We first need to calculate the sample proportionp. Out of 1,005 sampled adults, 643 knew what Viagra is used for, so[latex]\hat{p}=\frac{643}{1005}=.64[/latex]Therefore,A 95% confidence interval for p is[latex]\hat{\mathcal{p}}\pm2\ \bullet\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{\mathcal{n}}}=.64\pm2\sqrt{\frac{.64\left(1-.64\right)}{1005}}=.64\pm.03=\left(.61,\ .67\right)[/latex]We can be 95% sure that the proportion of all U.S. adults who were already familiar with Viagra by that time was between .61 and .67 (or 61% and 67%).The fact that the margin of error equals .03 says we can be 95% confident that unknown population proportion p is within .03 (3%) of the observed sample proportion .64 (64%). In other words, we are 95% confident that 64% is off by no more than 3%. We would like to share with you the methodology part of the poll release of the Viagra example, and show you that you now have the tools to understand how poll results are analyzed:The results are based on telephone interviews with a randomly selected national sample of 1,005 adults, 18 years and older, conducted May 8-10, 1998. For results based on samples of this size, one can say with 95 percent confidence that the error attributable to sampling and other random effects could be plus or minus 3 percentage points. In addition to sampling error, question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of public opinion polls. Two important results that we discussed at length when we talked about the confidence interval for also apply here:There is a trade-off between level of confidence and the width (or precision) of the confidence interval. The more precision you would like the confidence interval for p to have, the more you have to pay by having a lower level of confidence.Since n appears in the denominator of the margin of error of the confidence interval for p, for a fixed level of confidence, the larger the sample, the narrower, or more precise it is. This brings us naturally to our next point.Determining Sample Size for a Given Margin of Error for Estimating ProportionsJust as we did for means, when we have some level of flexibility in determining the sample size, we can set a desired margin of error for estimating the population proportion and find the sample size that will achieve that.For example, a final poll on the day before an election would want the margin of error to be quite small (with a high level of confidence) in order to be able to predict the election results with the most precision. This is particularly relevant when it is a close race between the candidates. The polling company needs to figure out how many eligible voters it needs to include in their sample in order to achieve that.Lets see how we do that.(Comment:For our discussion here we will focus on a 95% confidence level (z* = 2), since this is the most commonly used level of confidence.)The 95% confidence interval for p is[latex]\hat{\mathcal{p}}\pm\mathcal{z}*\ \bullet\sqrt{\frac{\mathcal{\hat{p}}\left(1-\hat{\mathcal{p}}\right)}{\mathcal{n}}}[/latex]The margin of error, then, is[latex]\mathcal{m}=2\sqrt{\frac{\hat{\mathcal{p}}\left(1-\hat{\mathcal{p}}\right)}{\mathcal{n}}}[/latex]Now we isolate n (i.e., express it as a function of m).[latex]\mathcal{n}=\frac{4\hat{\mathcal{p}}\left(1-\hat{\mathcal{p}}\right)}{\mathcal{m}^2}[/latex]There is a practical problem with this expression that we need to overcome.Practically, you first determine the sample size, then you choose a random sample of that size, and then use the collected data to find[latex]\hat{p}[/latex].So the fact that the expression above for determining the sample size depends on[latex]\hat{p}[/latex]is problematic.The way to overcome this problem is to take the conservative approach by setting [latex]\hat{p}=\frac{1}{2}[/latex].Why do we call this approach conservative?It is conservative because the expression that appears in the numerator,[latex]4\hat{p}(1-p)[/latex]is maximized when[latex]\hat{p}=\frac{1}{2}[/latex].That way, the n we get will work in giving us the desired margin of error regardless of what the value ofpis. This is a worst case scenario approach. So when we do that we get:[latex]\mathcal{n}=\frac{\left(4\right)\frac{1}{2}\left(1-\frac{1}{2}\right)}{\mathcal{m}^2}=\frac{1}{\mathcal{m}^2}[/latex]a conservative estimate for the margin of error, which is useful when we want to get a rough idea of its size without taking the trouble to make detailed calculations.Also, typically, there are several questions in a poll, each yielding a different [latex]\hat{p}[/latex]. Rather than reporting the separate margin of error for each question using [latex]\mathcal{m}=2\sqrt{\frac{\hat{\mathcal{p}}\left(1-\hat{\mathcal{p}}\right)}{\mathcal{n}}}[/latex], polls report just one, the conservative margin of error [latex]m=\frac{1}{\sqrt{n}}[/latex] as the margin of error of the poll, which is guaranteed to work for all the questions regardless of what the value of [latex]\hat{p}[/latex]ends up being. Aconfidence interval for a proportion is a range of values that is likely to contain a population proportion with a certain level of confidence.This tutorial explains the following:The motivation for creating a confidence interval for a proportion.The formula to create a confidence interval for a proportion.An example of how to calculate a confidence interval for a proportion.How to interpret a confidence interval for a proportion.Confidence Interval for a Proportion: MotivationThe reason to create a confidence interval for a proportion is to capture our uncertainty when estimating a population proportion.For example, suppose we want to estimate the proportion of people in a certain county that are in favor of a certain law. Since there are thousands of residents in the county, it would be too costly and time-consuming to go around and ask each resident about their stance on the law.Instead, we might select a simple random sample of residents and ask each one whether or not they support the law:Since we select a random sample of residents, there is no guarantee that the proportion of residents in the sample who are in favor of the law will exactly match the proportion of residents in the entire county who are in favor of the law.So, to capture this uncertainty we can create a confidence interval that contains a range of values that are likely to contain the true proportion of residents who are in favor of the law in the entire county.Confidence Interval for a Proportion: FormulaWe use the following formula to calculate a confidence interval for a population proportion:Confidence Interval = p +/- z*(p(1-p) / n)where:p: sample proportionz: the chosen z-valuen: sample sizeThe z-value that you will use is dependent on the confidence level that you choose. The following table shows the z-value that corresponds to popular confidence level choices:Confidence Levelz-value0.901.6450.951.960.992.58Notice that higher confidence levels correspond to larger z-values, which leads to wider confidence intervals.This means that, for example, a 95% confidence interval will be wider than a 90% confidence interval for the same set of data.Related: What is Considered a Good Confidence Interval?Confidence Interval for a Proportion: ExampleSuppose we want to estimate the proportion of residents in a county that are in favor of a certain law. We select a random sample of 100 residents and ask them about their stance on the law. Here are the results:Sample size = 100Proportion in favor of law = 0.56Here is how to find various confidence intervals for the population proportion:90% Confidence Interval:0.56 +/- 1.645*(.56(1-.56) / 100)=[0.478, 0.642]95% Confidence Interval:0.56 +/- 1.96*(.56(1-.56) / 100)=[0.463, 0.657]99% Confidence Interval:0.56 +/- 2.58*(.56(1-.56) / 100)=[0.432, 0.688]Note:You can also find these confidence intervals by using the Confidence Interval for a Proportion Calculator.Confidence Interval for a Proportion: InterpretationThe way we would interpret a confidence interval is as follows:There is a 95% chance that the confidence interval of [0.463, 0.657] contains the true population proportion of residents who are in favor of this certain law.Another way of saying the same thing is that there is only a 5% chance that the true population proportion lies outside of the 95% confidence interval.That is, theres only a 5% chance that the true proportion of residents in the county that support the law is less than 46.3% or greater than 65.7%.

**How to calculate confidence interval for proportion. How do you calculate a confidence interval for a proportion. Confidence interval for proportion.**