Continue

Two proportion z test example

This tutorial covers two-sample Z-tests for proportions, using both traditional and p-value approaches. The first example involves testing whether the percentage of women who use smartphones is less than the percentage of men who use smartphones. The null hypothesis assumes that both groups have the same proportion of smartphone users, while the alternative hypothesis suggests that women have a lower proportion of smartphone users. Using a significance level of 0.05 and the pooled estimate of sample proportions, the test statistic Z is calculated to be -2.476. Since this falls within the critical region (Z < -1.64), we reject the null hypothesis. Alternatively, using the p-value approach, the p-value for the observed test statistic is found to be 0.0066, which is less than the significance level of 0.05, leading us to reject the null hypothesis. The interpretation is that there is enough evidence to conclude that women have a lower proportion of smartphone users compared to men. The second example involves testing whether wearing seat belts reduces the fatality rate in car crashes. With a sample size of 2823 occupants not wearing seat belts and 7765 occupants wearing seat belts, we find that among the former group, 31 were killed, while among the latter group, 16 were killed. We test the claim that the fatality rate is higher for those not wearing seat belts using a significance level of 0.01. Let me know if you'd like me to rephrase anything further! **Problem Overview** A study compared the fatality rates of car occupants who wore seatbelts versus those who didn't. The results showed that 31 people died while not wearing seatbelts (out of 2823), and 16 people died while wearing seatbelts (out of 7765). The estimated sample proportions were calculated as 0.011 for those not wearing seatbelts and 0.002 for those wearing seatbelts. **Hypothesis Testing Problem** The research question is whether the fatality rate is higher for those not wearing seatbelts. The null hypothesis (H0) states that the two groups have equal fatality rates, while the alternative hypothesis (H1) states that the fatality rate is higher for those not wearing seatbelts. **Test Statistic** The test statistic Z was calculated using a formula involving the sample proportions and pooled estimate of sample proportion. The test statistic follows a standard normal distribution (N(0,1)). **Significance Level and Critical Value** The significance level α was set at 0.01, and the critical value for a right-tailed test was found to be 2.33 using a statistical table. **Computation and Decision** The test statistic Z_obs was calculated as 6.106, which falls within the rejection region (Z > 2.33). Alternatively, the p-value approach yielded a p-value of 0, which is less than the significance level α = 0.01. In both cases, the null hypothesis H0 was rejected. **Conclusion** There is sufficient evidence to support the claim that the fatality rate is higher for those not wearing seatbelts. This suggests that using seatbelts may be effective in saving lives. **Example Problem** The second part of the text presents an example problem where two machines are compared with respect to their proportion of defective parts. The sample proportions and pooled estimate of sample proportion were calculated, and a hypothesis testing problem was defined. **Similarities between Texts** Both texts involve hypothesis testing problems with a null and alternative hypothesis, test statistic calculation, significance level, critical value, computation, and decision-making process. The test statistic Z follows a standard normal distribution $N(0,1)$. ### Step 3: Specify the level of significance α $\alpha = 0.05$ ### Step 4: Determine the critical value Critical values for a two-tailed test are -1.96 and 1.96. ### Step 5: Computation $Z_{obs} = \frac{(0.118-0.255)-0}{\sqrt{\frac{0.195*(1-0.195)}{85}+\frac{0.195*(1-0.195)}{110}}} = -2.393$ ### Step 6: Decision Traditional approach: Since Z_{obs} falls inside the critical region, we reject the null hypothesis. p-value approach: The p-value is 0.0167, which is less than $\alpha = 0.05$, so we also reject the null hypothesis. ### Interpretation There is enough evidence to support the alternative hypothesis, suggesting that the two machines differ significantly with respect to the proportion of defectives. Given text here In statistics, probability sampling methods are used because every member of a population has an equal chance of being selected to be in the sample. The most common types of probability samples include simple random samples, stratified random samples, cluster random samples, and systematic random samples. In simple random samples, every member of a population is put into a hat and randomly drawn out, providing a representative sample. This method is useful when the researcher wants to know how a whole population feels about an issue. Stratified random samples divide a population into groups and then take a random sample from each group, which helps ensure that all subgroups are represented in the results. Cluster random samples group people together based on certain characteristics, such as schools or neighborhoods, and randomly select some of these groups to be included in the sample. Systematic random samples put all members in order and then choose every nth member to be in the sample. These types of samples are often used for research because they provide reliable data. Researchers can use different methods to collect data, including volunteer sampling, snowball sampling, and purposive sampling. Volunteer sampling involves asking participants to voluntarily participate in a study, which can lead to nonresponse bias and an unrepresentative sample. Snowball sampling involves recruiting initial subjects who then recruit additional subjects, which can result in sampling bias and a lack of representativeness. Purposive sampling involves selecting individuals based on specific criteria, which can also lead to a lack of representativeness. The two-proportion z-test is used to compare the proportions of two populations. The assumptions for this test are that both samples should be randomly drawn from their respective populations and follow a binomial distribution, with sample sizes greater than 10 and population sizes at least 10 times larger than the sample size. Hypotheses for the two-proportion z-test include a null hypothesis stating that the proportions of the two populations are equal, as well as alternative hypotheses that test whether one proportion is higher or lower than another. Ha : p1-p2 >0 The difference between two population proportions is greater than 0 i.e.proportion for population 1 is greater than the proportion for population 2.It is called Upper tail test (right-tailed test). Ha : p1-p2 ≠ 0 The difference between two population proportions is not equal to 0 i.e.proportion for population 1 is not equal to proportion for population 2.It is called two tail test. Formula for the test statistic two proportion Z test is: where : n1 : sample size for sample proportion from population 1. n2 : sample size for sample proportion from population 2. p1 : sample proportion for population 1 p2 : sample proportion for population 2 p : pooled sample proportion where To perform two proportion z-test, we will use the prop.test() functions from the R stats library. The prop.test() function uses the following basic syntax: prop.test(x, n, p = NULL, alternative = c("two.sided", "less", "greater"), conf.level = 0.95, correct = TRUE) where: x : Vector of number of successes n: Vector of a number of trials. p: The vector of probabilities of success. alternative: The alternative hypothesis for the test. It can be 'greater', 'less', 'two.sided' based on the alternative hypothesis. conf. level: confidence level of the interval correct: a logical indicating whether Yates' continuity correction should be applied or not Null Hypothesis: H0 : P1 = P2 (Population proportions for Coffee drinkers before and after excise duty are equal) Alternate Hypothesis: Ha: P1 > P2 (Population proportion for Coffee drinkers after excise duty is less than the Population proportion for Coffee drinkers before excise duty) Step 3: Calculate the test statistic using a prop.test() function in R: # Perform two-proportion z-test prop.test(x = c(800, 900), n = c(1000, 1200), alternative = "greater") Given text here is about performing a two-proportion z-test in R to determine if there's a significant difference between two population proportions. The test results show that the p-value (0.003115) is less than the level of significance (α = 0.05), indicating that we can reject the null hypothesis and conclude that the population proportions for coffee drinkers decrease after excise duty. To perform the two-proportion z-test in R, you'll need to use the "stats" package. The test results provide information about the X-squared value, degrees of freedom (df), p-value, alternative hypothesis, 95% confidence interval, and sample estimates. By interpreting these values, you can determine whether the observed difference between the two population proportions is statistically significant. The tutorial explains how to perform a two-proportion z-test in R, including the motivation behind using this test, the formula for performing it, and an example of how to do so. The test can be used to compare two population proportions, such as the proportion of residents who support a certain law in different counties. In summary, the two-proportion z-test is a useful statistical tool for comparing two population proportions and determining whether any observed differences are statistically significant. We use different types of hypotheses for the two-tailed, left-tailed, and right-tailed tests. The null hypothesis states that there is no difference between population proportions, while the alternative hypothesis states that there is a difference. For example, we can test if two population proportions are equal (H0: π1 = π2), not equal (H0: π1 ≠ π2), or one proportion is greater than the other (H0: π1 > π2) using the following formulas: z = (p1 - p2) / √(p * (1-p) * (1/n1 + 1/n2)) where: - z is the test statistic - p1 and p2 are sample proportions - n1 and n2 are sample sizes - p is the total pooled proportion If the p-value corresponding to the test statistic z is less than our chosen significance level, we can reject the null hypothesis. We will perform a two-proportion z-test using the following steps: Step 1: Gather the sample data. Sample 1: n1 = 50 p1 = 0.67 Sample 2: n2 = 50 p2 = 0.57 Step 2: Define the hypotheses. H0: π1 = π2 (the two population proportions are equal) H1: π1 ≠ π2 (the two population proportions are not equal) Step 3: Calculate the test statistic z. First, we will calculate the total pooled proportion: p = (p1n1 + p2n2)/(n1+n2) = (0.67(50) + 0.57(50))/(50+50) = 0.62 Next, we will calculate the test statistic z: z = (.67-.57) / √.62(1-0.62)(1/50+ 1/50) = 1.03 Step 4: Calculate the p-value of the test statistic z. According to the Z Score to P Value Calculator, the two-tailed p-value associated with z = 1.03 is 0.30301. Step 5: Draw a conclusion. Since this p-value is not less than our significance level α = 0.05, we fail to reject the null hypothesis. We do not have sufficient evidence to say that the proportion of residents who support this law is different between the two counties.