

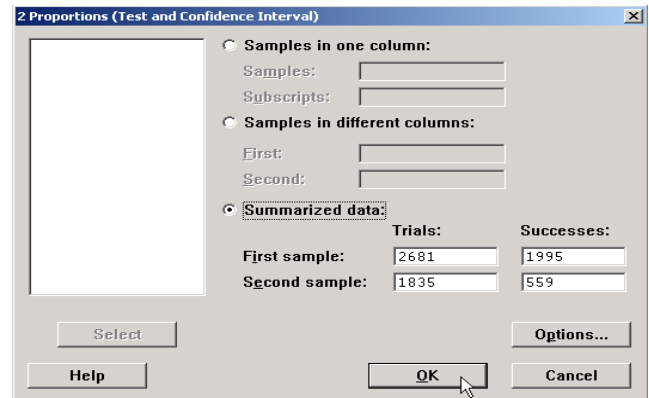
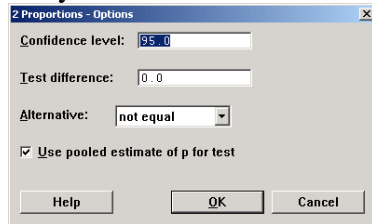
## Using Minitab

Choose **Stat** > **Basic Statistics** > **2 Proportions**

Enter the two sample sizes and the number of successes for each group.

Under the Options button, select “Use pooled estimate for test”

Remember to set the alternative to “not equal” to carry out the confidence interval



## Using Java applet

- Select two proportions option
- Specify direction of alternative and hypothesized difference
- Enter sample sizes then counts or proportions.

Two proportions

Ho:  $p_1 - p_2 = 0$

Ha:  $p_1 - p_2 < 0$

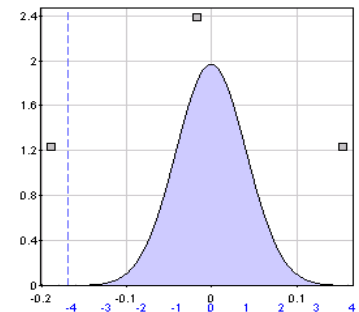
Group 1	Group 2
count: 13	count: 40
proportion: 0.08	proportion: 0.25
n: 164	n: 160

Calculate

test statistic:  $z = -4.13$

p-value: 0.0000

Reset



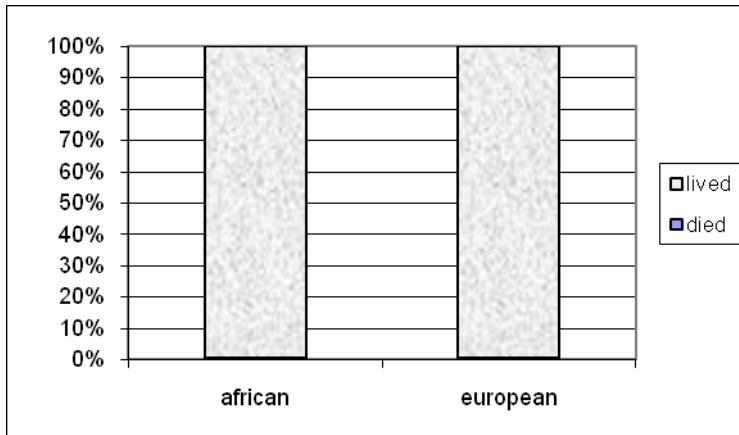
## Practice: The Gold Coast

A historian examining British colonial records for the Gold Coast in Africa suspects that the death rate was higher among African miners than among European miners. In the year 1936, there were 223 deaths among 33,809 African miners and 7 deaths among 1541 European miners on the Gold Coast. Is there good evidence that the proportion of African miners who died was higher than the proportion of European miners?

**Solution**

Sample 1 (African miners):  $n_1=33809$ ,  $\hat{p}_1 = 223/33809 = .006596$

Sample 2 (European miners):  $n_2 = 1541$ ,  $\hat{p}_2 = 7/1541 = .004543$



The death rates were very small but a slightly higher proportion of African miners died than European miners.

Let  $\pi_1$  = proportion of all African miners who died in the pre-war era in Africa

Let  $\pi_2$  = proportion of all European miners who died in the pre-war era in Africa

Hypotheses:

$H_0: \pi_1 - \pi_2 = 0$  (no difference in death rates)

$H_a: \pi_1 - \pi_2 > 0$  (higher proportion of African miners died)

*Technical conditions:*

At least 5 successes and failures in each group? Barely! Smallest was 7.

Random samples: Debatable...

From Minitab (Stat > Basic Statistics > 2 Proportions)

Sample	X	N	Sample p
1	223	33809	0.006596
2	7	1541	0.004543

Test for  $p(1) - p(2) = 0$  (vs  $> 0$ ):  $Z = 0.98$  P-Value = 0.163

*Decision:* Since this p-value is large we fail to reject  $H_0$ .

*Conclusion:* There is insufficient evidence to conclude that the proportion of African miners who died was higher than the proportion of European miners who died. However, we need to confirm that the death rates in 1936 are representative of the pre-war era in Africa.

**Summary:**

- Thoroughly check your technical conditions. Show the numbers if working with categorical data.
- It is important to remember all earlier lessons: start with graphical and numerical summaries (include them and discuss them), check technical conditions, make sure your hypothesis statements are about the parameter(s), interpret the p-value and confidence the same way, consider the study design when deciding if you can draw cause-and-effect conclusions or if you can generalize from the sample to a larger population.
- In comparing two proportions, make sure you are finding “the proportion of success” in both groups (as opposed to successes in one group and failures in the other for example). In dealing with conditional proportions, it is very important to keep straight which category is the one being conditioned on. Foreexample, the proportion of *American males* who are U.S. Senators is very small, yet the proportion of *U.S. Senators* who are American males is very large.
- I hope you will take care of the technology and let Minitab and/or the java applet do the calculations for you, allowing you to focus on selecting the appropriate procedure and interpreting the resulting output.