

2-Sample Categorical Data

BIOSTAT 201A Fall 2025

Discussion 7 – November 14, 2025

Cindy J. Pang

Contingency Tables	Two Sample Test for Difference in Population Proportions	Pearson's Chi-Square Test
Hypotheses		
Assumptions		
Data		
Test Statistic		
Decision Rule		

A hypothesis suggests that breast cancer risk increases with a longer interval between the onset of menstruation and a woman's first childbirth, making age at first birth a potential risk factor. To test this, an international study examined women from hospitals in the United States, Greece, Yugoslavia, Brazil, Taiwan, and Japan. Among women who had given birth, 21.2% of those with breast cancer (683 of 3,220) and 14.6% of those without breast cancer (1,498 of 10,245) had their first child at age 30 or older. The question is whether this difference reflects a true association or simply occurred by chance.

(a) Test this hypothesis using a 2-Sample Test for Difference in Population Proportions

A hypothesis suggests that breast cancer risk increases with a longer interval between the onset of menstruation and a woman's first childbirth, making age at first birth a potential risk factor. To test this, an international study examined women from hospitals in the United States, Greece, Yugoslavia, Brazil, Taiwan, and Japan. Among women who had given birth, 21.2% of those with breast cancer (683 of 3,220) and 14.6% of those without breast cancer (1,498 of 10,245) had their first child at age 30 or older. The question is whether this difference reflects a true association or simply occurred by chance.

(b) Test this hypothesis using a Pearson's Chi-square Test. *Are your conclusions different from part a?*