# **STAT 218 – Handout – Week 6 Lecture 1 – Difference of Two Proportions**

## **Part 1: Prenatal vitamins and Autism**

Researchers studying the link between prenatal vitamin use and autism surveyed the mothers of a random sample of children aged 24 - 60 months with autism and conducted another separate random sample for children with typical development. The table below shows the number of mothers in each group who did and did not use prenatal vitamins during the three months before pregnancy (periconceptional period).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Periconceptional prenatal vitamin* | | | |
| *Autism* |  | Vitamin | No Vitamin | Total |
| Autism | 143 | 111 | 254 |
| Typical Development | 159 | 70 | 229 |
| Total | 302 | 181 | 483 |

1. **Identify the observational unit *(singular, not plural)* and variables and classify the variable as quantitative or categorical.**

**A Mother of children aged 24-60 months is our observational unit. Variables are (1) whether a child has autism or typical development – categorical binary and (2) whether mother took periconceptional prenatal vitamin or not – categorical binary.**

1. **Which would you consider the explanatory variable in this study? Which is the response?**

**Developing Autism is the response variable while taking prenatal vitamins is the explanatory variable.**

1. **Describe in words the relevant parameter of interest in this study. What symbol would you use to represent it?**

**Difference in population proportions is our parameter which is π1 – π2. Difference in the rates of autism of children of mothers who did and did not use prenatal vitamins during the first three months before pregnancy.**

1. **Calculate the relevant statistic in this study. What symbol would you use to represent it?**

**To be able to do that, we need conditional proportions (column percentages)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Periconceptional prenatal vitamin* | | | |
| *Autism* |  | Vitamin | No Vitamin | Total |
| Autism | 143 **(47.4%)** | 111  **(61.3%)** | 254 |
| Typical Development | 159  **(52.6%)** | 70  **(38.7%)** | 229 |
| Total | 302  **(100.0%)** | 181  **(100%)** | 483 |

**Our sample statistic is 1 – 2  = 0. 474 – 0.613 = -0.139**

1. **State appropriate null and alternative hypotheses to test the association between use of prenatal vitamins during the three months before pregnancy and autism.**

**H0: π1 – π2 = 0 (There is no difference in the rates of autism of children of mothers who did and did not use prenatal vitamins during the first three months before pregnancy.) HA: π1 – π2 ≠ 0 (There is some difference in the rates of autism of children of mothers who did and did not use prenatal vitamins during the first three months before pregnancy.)**

1. **Let’s use** [**simulation-based approach**](https://www.rossmanchance.com/applets/2021/chisqshuffle/ChiSqShuffle.htm?FET=1) **to see if there is any association between use of prenatal vitamins during the three months before pregnancy and autism. Report your simulation-based p value.**

**A screenshot of a computer

AI-generated content may be incorrect.**

Our simulation-based p-value is 0.004 so we reject the null hypothesis. We have very strong evidence against the null hypothesis that there is no difference in the rates of autism of children of mothers who did and did not use prenatal vitamins during the first three months before pregnancy.

**Interpretation of p-value:** Under the assumption that the usage of prenatal vitamin does not make a difference, if we repeated the random assignment process many, many times, we would get results as extreme or more extreme than 0.1397, in about 0.4% of the repetitions.

1. **Can we use theory-based inference in this example? Verify any necessary validity conditions for the theory-based inference.**

**We have at least 10 successes and 10 failures in each group, so we satisfied the conditions.**

1. **Use Theory-based inference Applet to find standardized statistic, standard error, theory-based p-value and 95% confidence interval for the parameter π1 – π2. Report them below.**

**Standard error: 0.047; Standardized statistic: -2.98 =**

**Theory-based p-value: 0.0029 (again very strong evidence against the null hypothesis)**

**A screenshot of a computer

AI-generated content may be incorrect.95%CI = (-0.2303, -0.0492)**

**REMEMBER!** When the validity conditions are satisfied, the simulation-based and theory-based approaches will produce very similar results and lead to the same conclusions.

That’s because:

* The **Central Limit Theorem** ensures the sampling distribution is approximately normal.
* Simulation-based methods empirically approximate that same distribution.

However, if the validity conditions are not met, the simulation-based approach can be reported because simulation-based methods are more robust when theoretical assumptions fail (e.g., small samples, skewed data).

1. **Interpret hypothesis testing results and confidence interval and state an appropriate conclusion.**

**We have very strong evidence against the null hypothesis of no association between taking prenatal vitamins and autism. We are 95% confident that the rate of autism among children whose mothers took prenatal vitamins before pregnancy is between 4.9% and 23.0% lower than among children whose mothers did not take prenatal vitamins.**

1. **New York Times article reporting on this study was titled “Prenatal Vitamins May Ward Off Autism". Do you find the title of this article to be appropriate? Explain your answer.**

**The title of this newspaper article makes it sound like using prenatal vitamins can prevent autism, which is a causal statement. Since this is an observational study, we cannot make causal statements based on the findings of the study.**

1. **Propose an alternative title.**

**A more accurate title would be Mothers who use prenatal vitamins before pregnancy are found to have children with a lower rate of autism.**

## **Part 2: Prenatal vitamins and Autism - Relative Risk**

|  |
| --- |
| **Definition:** ***Relative risk*** is the ratio of two conditional proportions. It indicates how many times greater the risk of an outcome is for one group compared to the risk for the other group. |

1. **Calculate the relative risk of developing autism by dividing the proportion of autism cases among children whose mothers did not take prenatal vitamins before pregnancy by the proportion of autism cases among children whose mothers took prenatal vitamins.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Periconceptional prenatal vitamin* | | | |
| *Autism* |  | Vitamin | No Vitamin | Total |
| Autism | 143  **(47.4%)** | 111  **(61.3%)** | 254 |
| Typical Development | 159  **(52.6%)** | 70  **(38.7%)** | 229 |
| Total | 302  **(100.0%)** | 181  **(100%)** | 483 |

**0.613/0. 474 = 1.29**

1. **Write a sentence interpreting this ratio value. Does the value of this ratio strike you as noteworthy?**

**The relative risk of developing autism is 1.29, meaning that children whose mothers did not take prenatal vitamins before pregnancy are 1.29 times as likely to develop autism as those whose mothers did take prenatal vitamins.**

**Because the relative risk is greater than 1, this indicates a higher risk of autism for children of mothers who did not take prenatal vitamins before pregnancy.**