

Project #2: Probability Summary

Recall the data on birth weight gathered from the North Carolina State Center for Health and Environmental Statistics. Use the data set of 1450 observations **ncbirth650.MTW**. Recall the variables the variables examined are:

Variable Label	Description
PLURALITY	Number of children born of the pregnancy
SEX	Sex of child (1=Male, 2=Female)
MAGE	Age of mother (years)
WEEKS	Completed Weeks of Gestation (weeks)
MARITAL	Marital status (1=married, 2=not married)
RACEMOM	Race of Mother (0=Other Non-white, 1=White, 2=Black 3=American Indian, 4=Chinese, 5=Japanese, 6=Hawaiian, 7=Filipino, 8=Other Asian or Pacific Islander)
HISPMOM	Mother of Hispanic origin (C=Cuban, M=Mexican, N=Non-Hispanic, O=Other and Unknown Hispanic, P=Puerto Rican, S=Central/South American, U=Not Classifiable)
GAINED	Weight gained during pregnancy (pounds)
SMOKE	0=mother did not smoke during pregnancy 1=mother did smoke during pregnancy
DRINK	0=mother did not consume alcohol during pregnancy 1=mother did consume alcohol during pregnancy
TOUNCES	Weight of child (ounces)
TGRAMS	Weight of child (grams)
LOW	0=infant was not low birth weight 1=infant was low birth weight
PREMIE	0=infant was not premature 1=infant was premature premature defined at 36 weeks or sooner

Pregnant women are encouraged by physicians to refrain from such activities as smoking. It is believed that women who engage in these activities are at higher risk for having a low birth weight child. This assignment will have you investigate whether this actually turns out to be true with the data above.

Create a contingency table of **smoke** vs. **low**. Give an estimate of the probability of a low birth weight child. Give an estimate of the probability that the child's mother smokes. Estimate the probability of low birth weight given the mother smokes. Estimate the probability of low birth weight given the mother does not smoke. Calculate the prevalence ratio of low birth weight vs. smoking status and interpret its meaning. (Remember to have MINITAB ignore all missing values).

Obviously, premature babies are more likely to be low birth weight, but how obvious? Create a table of **premie** vs. **low**. Calculate the probability a baby is born premature. Calculate the probability the baby is low birth weight given it is premature. Calculate the probability the baby is low birth weight given it is not premature. Calculate the prevalence ratio of low birth weight if the baby is premature. Interpret its meaning.

Create a table of **smoke** vs. **premie**. Using **smoke** as the risk factor, calculate the probability of a premature baby given the mother smoked, the probability of a premature baby given the mother did not smoke, and the prevalence ratio of a premature baby with smoking as the risk factor. Interpret your results.

Create a table of **marital** vs. **low** and **marital** vs. **premie**. Using **low** as the outcome variable, calculate the prevalence ratio comparing the probability of **low** for unmarried mothers to married mothers. Do the same for the outcome variable of **premie**. Interpret your results.

Dr. Holcomb's mother is quite angry at the above analysis. She smoked during her two pregnancies and both of her children weighed above 6 pounds (ie they were not low birth weight). She thinks her son has wasted his life studying a useless subject like statistics because she smoked and her kids turned out fine. Write a paragraph explaining the results of this analysis and how they relate to Mrs. Holcomb's experience. Explain to Mrs. Holcomb why the above analysis points to a recommendation curtailing smoking during pregnancy.

After doing the analysis above, a MTH 147 student proclaims that being unmarried causes premature birth and low birth weight. Is this logic correct? Explain in words why or why not. (Hint: Give a confounding variable and explain how it is related to both the explanatory and response variables).