Statistics 200B
Spring 1997

Lab V: Babies and Smoke

1 Introduction

The American Cancer Society billboards and television commercials make very strong statements against pregnant women smoking cigarettes. A Surgeon General's Warning found on the side panel of packages of cigarettes reads:

> Smoking by pregnant women may result in fetal injury, premature birth, and low birth weight.

In this laboratory, you will have the opportunity to analyze a study of pregnant women to see whether the Surgeon General's Warning is corroborated. The data provided here is part of the Child Health and Development Studies (CHDS) – a comprehensive investigation of all pregnancies that occurred between 1960-1967 among women in the Kaiser Foundation Health Plan in the San Francisco-East Bay area. This study is notable because it found that babies of smokers did not have an increased neonatal mortality in comparison to nonsmokers.

Despite the warnings of the Surgeon General, the American Cancer Society and health care practitioners, pregnant women do smoke. The 1985 National Health Interview Survey [1] found that 25% of women who were pregnant in 1980-85 smoked during their pregnancy.

Unlike the case for smoking and lung cancer, the word *may* in the warning above reflects the fact that research on the effects of smoking during pregnancy are not yet conclusive. According to the 1989 Surgeon General's Report [3],

> ... cigarette smoking seems to be a more significant determinant of birthweight than the mother's prepregnancy height, weight, parity, payment status, or history of previous pregnancy outcome, or the infant's sex. The reduction in birthweight associated with

> maternal tobacco use seems to be a direct effect of smoking on fetal growth.
>
> Mothers who smoke also have increased rates of premature delivery. The newborns are also smaller at every gestational age. The infants display symmetrical fetal growth retardation with the deficits in measurements of crown-heal length, chest and head circumferences, and birthweight.

Other studies [7] indicate that smoking is responsible for a 150 to 250 gram reduction in birth weight, and that smoking mothers are about twice as likely as nonsmoking mothers to have a low birth weight baby (under 2,500 grams). Birth weight is a measure of the baby's maturity, as is the baby's gestational age, or the time spent in utero. Smaller babies and preterm babies have much higher rates of infant mortality than larger babies who are born at term.

2 The Data

The data for this lab assignment is a subset of a much larger study – the Child Health and Development Study [4]. The CHDS database includes all pregnancies that occurred between 1960-1967, among women in the Kaiser Foundation Health Plan in Oakland California. The Kasier Health Plan is a prepaid medical care program. The members represent a broadly based group which is typical of an employed population. Approximately 15,000 families participated in the study.

The women were interviewed early in their pregnancy. A variety of medical, genetic, and environmental information was ascertained on both parents, including height, weight, age, smoking habits, and education. At birth, measurements on the baby were recorded. They included the baby's length, weight, and head circumference. Ten years later, within two weeks of the child's birthday, the child and mother were again called for more measurements. At this time, measurements such as the height and weight of the child were taken as well as scores on verbal and math tests.

Provided here, is a subset of the information collected for 1236 boy babies. The variables that are available for analysis are described in the table below. The TA will give you instructions about how to access the data.

```
1 idnumber
2 pluralty=5 for single fetus
3 outcome=1 for live birth that survived at least 28 days
4 birth date 1096=January1,1961
5 gestation days
6 infant sex 1=male 2=female 9=unknown
7 birth weight in ounces (999 unknown)
8 parity 0=first born
9 mom race 0-5=white 6=mex 7=black 8=asian 9=mix 99=unknown
10 mom age in years
11 mom education 0= (<8), 1=(8-<12), 2=12, 3=12+trade, 4=12+some college
              5=16, 6,7 trade (hs unclear), 9=unknown
12 mom height in inches
13 mom prepregnancy wt in pounds
14 dad race
15 dad age
16 dad educ
17 dad height
18 dad weight
19 marital 1=married, 2-4=sep.div,wid, 5=never married, blank
20 total income in 2500 increments 0=under 2500, 1=2500-4999, ..., 9=15000+
   98=unknown, 99=not asked
21 mom smoke 0=never, 1=yes,now, 2=until preg, 3=once did, not now,
    9=unknown
22 how long ago quit 0=never, 1=still, 2=during preg, 3=up to 1 yr,
   4=up to 2 yr, 5=up to 3 yr, 6=up to 4 yr, 7=5 to 9yr, 8=10+yr, 9=quit
   and don't know, 98=unknown
23 number of cigs smoked a day for past and current smokers  0=never,
   1=1-4 2=5-9, 3=10-14, 4=15-19, 5=20-29, 6=30-39, 7=40-60, 8=60+,
   9=smoke but don't know, 98=unknown
```

3 Background

The typical gestation period for a baby is 40 weeks [5]. Those born earlier than 37 weeks are considered preterm. Few babies are allowed to remain in utero for more than 42 weeks, for brain damage may occur due to deterioration of the placenta. At 28 weeks of age, the fetus weighs about 4 to 5 pounds and is about 16 inches long. At 32 weeks, it weighs about 5 to 5.5 pounds and is about 18 inches long. In the immediate weeks prior to delivery, babies gain about 2/10 of a pound a week. Most newborns range from 18 to 22 inches in length and from 5.5 to 8.8 pounds (1 pound is 454 grams). Babies born at term that weigh less that 5.5 pounds are considered small for their gestational age.

Before the 1940s it was widely believed that the baby was in a protected state while in the uterus. This theory was attacked in the 1940s when a pediatrician from Melbourne Australia found a connection between babies born to mothers who contracted rubella in an epidemic in their first trimester and heart, eye and ear defects. One infamous example of the effects of chemical agents on the the unborn baby is the thalidimide tragedy of the 1960s. Several thousand babies were born with limb defects as a result of their mothers taking thalidimide early in pregnancy.

It is commonly thought [2] that smoking reduces the oxygen supplied to the fetus. The physiological explanation for this reduction is that the carbon monoxide in the smoke the mother inhales binds with the hemoglobin in the mother's blood to form carboxyhemoglobin. Hemoglobin has a 210 times greater affinity for carbon monoxide than oxygen. The increased levels of carboxyhemoglobin restrict the amount of oxygen that can be carried by the blood. For the fetus, this problem is compounded when the oxygen is released from the mother's blood to the placenta to be taken up again by the fetus' blood. A steady supply of oxygen is critical for the developing baby. It is hypothesized, that to compensate for the decreased supply of oxygen, the placenta increases in size, surface area, and vascularity, and the fetus increases the levels of hemoglobin and redistributes the blood flow to favor vital parts. These same survival mechanisms are observed in high altitude pregnancies.

4 Investigations

What is the effect of the mother's smoking on the fetus? The questions below are meant to help guide you in answering this question. The aim of the analysis is exploratory, using primarily graphical techniques to reveal possible trends and relationships.

Begin by comparing the distributions of physical measurements of the babies of heavy smokers, light smokers, and nonsmokers. To do this split the data into three groups corresponding to the number of cigarettes per day smoked by the mother: (1) nonsmoker, (2) moderate smoker (less than 20 per day), (3) heavy smoker (more than 20 per day).

- What is the mean and median birth weight for each of the three sets of babies? What is the standard deviation?
- Make boxplots that compare the distributions of birth weight for the three groups.
- Make a similar comparison using histograms or smoothed density functions. In either case you will have to experiment with the amount of smoothing—the bin width for the histogram or the bandwidth for the density estimate. A rough guideline that is often effective in choosing the bandwidth for a density function estimate based on a Gaussian window is to set the standard deviation of the Gaussian window equal to $\sigma n^{-1/5}$, where n is the number of data points and σ is their standard deviation.
- Another technique for comparing the distributions is to plot the percentiles of weight for one group against the percentiles for weight of the other group (see Section 10.2.3 of the text). Do this for the three pairs.

Summarize your findings about differences in the distributions of birth weight between these three groups in one or two paragraphs, referring to the results found above as necessary. Which techniques did you find helpful in revealing differences?

The data we are analyzing is from an observational study, not a designed experiment. The families that participated in the study chose which group they belonged to, because the mother decided whether or not to smoke during her pregnancy. Smokers may differ from nonsmokers in some very essential ways. That is, they may differ in ways that will effect the birthweight of the baby, whether or not the mother smoked. It is in general difficult to defend causal inferences made from observational studies. According to Malloy et al [6], in 1985, pregnant women in the state of Missouri that smoked tended to be less well educated and younger than those that did not smoke; they also tended to have more children. According to Yerushalmy [4], the nonsmokers in the CHDS study were more likely to use contraceptive methods, to plan the baby, and less likely to drink coffee and hard liquor. In what other ways that could affect birthweight would you imagine that the smokers and nonsmokers differ?

- Make comparisons as you did for the babies weight, but this time compare characteristics of the mothers and fathers for smokers versus nonsmokers. Start by looking at three factors: mother's prepregnancy weight, father's education, and mother's height. Are there any differences of note between the two groups with respect to these factors?

- Examine the relationships of the factors you identify to birthweight. For example, make a scatterplot of birthweight versus mother's weight, identifying points that correspond to smokers and nonsmokers. Interpret what you see.

- Is it possible that any differences in birthweight that you see are due to differing gestational ages for smokers and nonsmokers? Make a scatterplot of birth weight and gestational age identifying points corresponding to smokers and nonsmokers. What do you see?

5 References

[1] National Center for Health Statistics (1988) Health promotion and disease prevention: United States, 1985. *Vital and Health Statistics*, Series 10, No. 163, Public Health Service. DHHS Publication No. (PHS) 88-1591, February 1988.
[2] Longo, L. (1976) Carbon monoxide: Effects on oxygenation of the fetus in utero. *Science* vol. 194, p 523.
[3] United States Department of Health and Human Services (1989) Reducing the health consequences of smoking: 25 years of progress. A report of the Surgeon General. *US Department of Health and Human Services, Public Health Service, Office on Smoking and Health.* DHHS publication No. (CDC) 89-8411, 1989.
[4] Yerushalmy, J. (1971) The relationship of parents' cigarette smoking to outcome of pregnancy – Implications as to the problem of inferring causation from observed associations. *Amer J Epid* vol 93, pp 443-456.

[5] Samuels, M. & Samuels, N. (1986) *The Well Pregnancy Book.* Summit Books, New York.
[6] Malloy, M., Kleinman, J., Land, G., Schram, W. (1988) The association of maternal smoking with age and cause of infant death. *Amer J Epid* vol 128, pp 46-55.
[7] Merkatz, I. & Thompson, J. (1990) *New Perspectives on Prenatal Care.* Elsevier, New York.