

STAT 1020: Discussion Handout 1

Chapter 1 Terminology

(Adapted from question 25, textbook)

Babies Medical researchers at a large city hospital investigating the impact of prenatal care on newborn health collected data from 882 births during 1998–2000. They kept track of the mother's age, the number of weeks the pregnancy lasted, the type of birth (cesarean, induced, natural), the level of prenatal care the mother had (none, minimal, adequate), the birth weight and sex of the baby, and whether the baby exhibited health problems (none, minor, major).

- a) What is the sample and population of this study?

Sample: The 882 births

Population: All births

- b) List the who, what, when, where, why, and how of the study. If no information is given on a particular W, write "NA".

Who: The 882 births

What: Mother's age, duration of pregnancy, type of birth, level of prenatal care of mother, birthweight of baby, sex of baby, presence of health problems

When: 1998 – 2000

Where: NA (While the hospital seems like a likely location, we only know the researchers are from there, not the babies)

Why: To investigate the impact of prenatal care on newborn health

How: NA (although we can guess)

- c) For each of these quantities, state if they are categorical variables (C), quantitative variables (Q), or not a variable (N).

Number of births: N

Birth weight of baby: Q

Mother's age: Q

Sex of baby: C

Duration of study: N

Level of prenatal care: C

Duration of pregnancy: Q

Average birth weight: N

Type of birth: C

Most common level of prenatal care: N

Whether the baby exhibited health problems: C

Average duration of pregnancy: N

Chapter 2: Contingency table and conditional distributions

(adapted from Statcrunch data set)

Forecast\Actual	No Rain	Rain	Total
No Rain	268	7	275
Rain	63	27	90
Total	331	34	365

The above is a contingency table of actual weather data against the corresponding forecasted weather for a year.

a) What percent of days in the year was there (actually) no rain?

$$\begin{aligned}\% \text{ of days actually no rain} &= \frac{\text{Total \# of days with (actual) no rain}}{\text{Total \# of days in year}} \\ &= (268 + 63)/365 \times 100\% = 90.68\%\end{aligned}$$

b) What percent of days in the year was there no rain, and the forecast matched the actual weather?

$$\begin{aligned}\% \text{ of days forecast and actual "no rain"} &= \frac{\text{\# of days actual no rain, forecast no rain}}{\text{Total \# of days in year}} \\ &= 268/365 \times 100\% = 73.42\%\end{aligned}$$

c) What percent of days when there was no rain, was the forecast accurate?

$$\begin{aligned}\% \text{ of days forecast no rain out of actual no rain days} &= \frac{\text{\# of days actual no rain, forecast no rain}}{\text{Total \# of days (actual)no rain}} \\ &= 268/331 \times 100\% = 80.97\%\end{aligned}$$

d) What percent of days when rain was forecasted, was the forecast accurate?

% of days actual rain out of days when rain was forecasted

$$\begin{aligned}&= \frac{\text{\# of days actual rain, forecast rain}}{\text{Total \# of days (forecast) rain}} \\ &= 27/90 \times 100\% = 30\%\end{aligned}$$

- e) Fill in the blanks in the contingency table with the **relative frequency, conditioned on the actual outcome of the weather.** (Extra practice after class: also make sure you know how to fill it in, if conditioned on the forecasts of the weather.)

If conditioned on actual outcome of weather:

Forecast\Actual	No Rain	Rain	Total
No Rain	268 (80.97 %)	7 (20.59 %)	275 (75.34 %)
Rain	63 (19.03 %)	27 (79.41 %)	90 (24.66 %)
Total	331 (100 %)	34 (100 %)	365 (100 %)