TYPES OF SAMPLING AND STUDIES

Sections 2.1.4 and 2.3.6

Types of Sampling in Contingency Tables: Overview

1. Poisson
2. Binomial
3. Multinomial

Name denotes the probability distribution that the cell counts in a contingency table follow (before data collection).

Poisson Sampling

- The sample size (n) is <u>random</u>, as are all the cell counts (n_{ij})
- Each cell count assumed to have a Poisson distribution

Example:

Interview every 5th customer exiting a restaurant for one hour and ask X = gross annual income level (Under $25k, $25k-$50k, etc) and Y = answer to "how satisfied were you by your meal?" (Very dissatisfied, Dissatisfied, Neither satisfied nor dissatisfied, Satisfied, Very satisfied)

Multinomial Sampling

- The sample size (n) fixed before data collection, but the cell counts and marginal totals are random
- Take a sample of size n and cross-classify into two categorical variables
- Vector of cell counts ($n_{11}, n_{12}, \ldots, n_{IJ}$) follows a multinomial distribution

Example:

Select a simple random sample of 100 adults, then classify according to X = who they voted for in the last presidential election, and Y = who they plan to vote for in the next presidential election

Binomial Sampling

- The sample size (n) and the row (column) totals are fixed before data collection, but the cell counts and column (row) totals are random
- Vector of cell counts within the ith row (jth column) follows a binomial (if two possible categories) or multinomial (if more than two categories) distribution

Example:

Select a random sample of 500 smokers and 500 non-smokers, then follow for 20 years and determine Y = whether they contracted lung cancer

What can we estimate?

- Poisson sampling → joint, conditional, and marginal probabilities
- Multinomial sampling → joint, conditional, and marginal probabilities
- Binomial sampling → only conditional probabilities (within a category whose sample size was fixed) and marginal probabilities (of the variable whose sample sizes were not fixed)

Types of Studies: Overview

- Prospective vs. retrospective studies
- Observational studies vs. randomized experiments (experimental studies)
  - Observational studies:
    - Cohort (prospective)
    - Case-control (retrospective)
    - Cross-sectional
  - Experimental studies:
    - Clinical trial (prospective)

Observational Study vs. Experiment

- An **observational study** is one in which participants are merely observed or surveyed and measured.
  - Comparisons based on observational studies are comparisons of naturally occurring groups.
  - ***Cause and effect conclusions generally cannot be made from an observational study. Why?***

Confounding Variables

- A **confounding variable** is a variable that both *affects the response variable* and also is *related to the explanatory variable*.
  - The effect of a confounding variable on the response variable cannot be separated from (is "confounded" with) the effect of the explanatory variable.
- Example *USA Today* headline: "**Prayer can lead to lower blood pressure**"
  - Reports an observational study reported in which those who attended religious services had lower blood pressure than those who tuned into religious TV or radio. Possible confounding variables?

Observational Study vs. Experiment

- In a **randomized experiment** participants are randomly assigned to treatment groups (explanatory variable).
  - The random assignment should balance out values of confounding variables across groups.
  - With a well-designed experiment, cause and effect conclusions generally *can* be made.

Prospective vs. Retrospective

- A **retrospective** design collects the sample, then asks about the past:
- Example:
  - **Case-control study** = *observational study* in which a sample of "cases" is collected and a sample of "controls" is collected (based on the outcome variable); then compare explanatory variables of interest

Prospective vs. Retrospective

- A **prospective** design collects the sample, then follows the individuals for a length of time and observes future events during the study period:
- Examples:
  - **Cohort study** = *observational study* in which a cohort (group of people who share a common characteristic) is followed throughout the study and outcome variables are observed
  - **Clinical trial** = *randomized experiment* in which individuals are randomly assigned to treatments then followed throughout the study and outcome variables are observed

Prospective vs. Retrospective

• In a **cross-sectional study**, a representative sample of a population is collected, and data are collected on those individuals at a specific point in time.
  • A cross-sectional study is an observational study that is neither prospective or retrospective.

Example 1

Sample of 42 UC-Irvine intro statistics students from fall quarter 2012 that responded to an online survey; classified by whether or not they commuted to school and whether or not they were "usually late for class".

	Usually late for class	Not usually late for class	TOTAL
Commuted	8	4	12
Non-commuter	6	24	30
TOTAL	14	28	42

Type of sampling?

Type of study?

Example 2

Researchers at Mayo Clinics recruited 240 smokers with a desire to quit. Half of the volunteers were randomly assigned to receive a nicotine patch, and the other half received a placebo patch, each for 8 weeks. After the 8-week period, researchers recorded if they quit or did not quit smoking (*Journal of American Medicine*, 1994).

	Quit	Did not quit	TOTAL
Nicotine	56	64	120
Placebo	24	96	120
TOTAL	80	160	240

Type of sampling?

Type of study?

Example 3

Researchers at Boston University School of Medicine compared 665 men who had been admitted to the hospital with their first heart attack to 772 men in the same age group (21 to 54 years old) who had been admitted to the same hospitals for other reasons. They measured the degree of male-pattern baldness for each patient (*Journal of the American Medical Association*, 1993).

	Heart Attack (cases)	No Heart Attack (controls)	TOTAL
Baldness	279	263	542
No baldness	386	509	895
TOTAL	665	772	1437

Type of sampling?

Type of study?

Example 4

Medical researchers followed 6272 Swedish men for 30 years to see if there was an association between the amount of fish in their diet and prostate cancer ("Fatty Fish Consumption and Risk of Prostate Cancer," *Lancet*, June 2001).

		Prostate Cancer		TOTAL
		No	Yes	
Fish Consumption	Never/seldom	110	14	124
	Small part of diet	2420	201	2621
	Moderate part	2769	209	2978
	Large part	507	42	549
	TOTAL	5806	466	6272

Type of sampling?

Type of study?