Chapter 2: Samples, Good and Bad?

Thought Question: Write-in interview

Popular magazines often contain surveys that ask their readers to answer questions about hot topics in the news.

*Do you think the responses the magazines receive are representative of public opinion?*

Thought Question: Call-in interview

Cable News Network (CNN) often asks its viewers to call the network with their opinions on certain political issues, like whether or not they favor current foreign policy.

(i) *Do you think the results of these polls represent the feelings of the general population?*

(ii) *Do you think they represent the feelings of all those watching CNN at the time?*

Thought Question: Interview at the mall

Manufacturers and advertising agencies often use interviews at shopping malls to gather information about the habits of consumers and the effectiveness of ads. A sample of mall shoppers is fast and cheap.

*Do you think the results represent the whole U.S. population?*
Bad Ways to Sample

Terminology

- **Convenience Sampling**: selecting individuals who are easiest to reach.
- **Voluntary Response Sampling**: allowing individuals to choose to be in the sample.
- **Biased Sampling Methods**: systematically favor certain outcomes.

Both convenience sampling and voluntary response sampling are biased.

Example: Convenience Sampling

A research group wants to sample mice from a large cage to study how a drug affects physical activity. The lab assistant reaches into the cage to select the mice one at a time until 10 are chosen.

*Which mice will likely be chosen?* the largest ones

*Could this sample yield biased results?* yes

Example: Voluntary Response Sampling

Ann Landers once asked the readers of her advice column, “If you had it to do over again, would you have children?” Out of nearly 10,000 responses almost 70% said “NO!”.

*Do you think it’s true that 70% of parents regret having children?* no

*What’s actually going on?*

It is most likely that people feeling strongly about an issue will take the trouble to respond.
**Example: 1936 Presidential Election**

The 1936 presidential election proved to shape the future of opinion polling. The Literary Digest, the venerable magazine founded in 1890, had correctly predicted the outcomes of the 1916, 1920, 1924, 1928, and 1932 elections by conducting polls. These polls were a lucrative venture for the magazine: readers liked them; newspapers played them up; and each “ballot” included a subscription blank. The 1936 postal card poll claimed to have asked one forth of the nation’s voters which candidate they intended to vote for. In Literary Digest’s October 31 issue, based on more than 2,000,000 returned post cards, it issued its prediction: “Republican candidate Alfred Landon would win 57% of the popular vote.” In the election, FDR won 62 percent of the popular vote in one of the biggest landslides in election history!

The failure of the Literary Digest poll can be attributed to bad sampling techniques.

- The sample itself was a voluntary response sample. Some people who received questionnaires responded, but most did not.

- The sampling frame was biased; the mailings went to people who had 1936 auto registrations, who were listed in telephone books, and who were on Literary Digest’s subscription list!

Literary Digest went bankrupt soon after this debacle. Meanwhile, George Gallup was beginning to use random samples and scientific methods, and the survey research industry was born.
**Good Way to Sample: Simple Random Sampling (SRS)**

**Terminology**

- **Simple Random Sample:**
  - Each individual in the population has the same chance of being chosen for the sample.
  - Each group of individuals in the population of the required size \( n \) has the same chance of being the sample actually selected.

**Ways to do an SRS**

- drawing names out of a hat

- random number table

- computer software

**Example: SRS**

John’s small accounting firm serves 30 business clients. John wants to interview a sample of 5 clients to find ways to improve client satisfaction. To avoid bias, he chooses an SRS of size 5.

**Step 1:** *Give each client a numerical label, using as few digits as possible.*
<table>
<thead>
<tr>
<th>Label</th>
<th>Client</th>
<th>Label</th>
<th>Client</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>A-1 Plumbing</td>
<td>16</td>
<td>JL Records</td>
</tr>
<tr>
<td>02</td>
<td>Accent publishing</td>
<td>17</td>
<td>Johnson Commodities</td>
</tr>
<tr>
<td>03</td>
<td>Action Sport Shop</td>
<td>18</td>
<td>Keiser Construction</td>
</tr>
<tr>
<td>04</td>
<td>Anderson Construction</td>
<td>19</td>
<td>Liu’s Chinese Restaurant</td>
</tr>
<tr>
<td>05</td>
<td>Bailey Trucking</td>
<td>20</td>
<td>MagicTan</td>
</tr>
<tr>
<td>06</td>
<td>Balloons, Inc.</td>
<td>21</td>
<td>Peerless Machine</td>
</tr>
<tr>
<td>07</td>
<td>Bennett Hardware</td>
<td>22</td>
<td>Photo Arts</td>
</tr>
<tr>
<td>08</td>
<td>Best’s Camera Shop</td>
<td>23</td>
<td>River City Books</td>
</tr>
<tr>
<td>09</td>
<td>Blue print specialties</td>
<td>24</td>
<td>Riverside Tavern</td>
</tr>
<tr>
<td>10</td>
<td>Central Tree Service</td>
<td>25</td>
<td>Rustic Boutique</td>
</tr>
<tr>
<td>11</td>
<td>Classic Flowers</td>
<td>26</td>
<td>Satellite Services</td>
</tr>
<tr>
<td>12</td>
<td>Computer Answers</td>
<td>27</td>
<td>Scotch Wash</td>
</tr>
<tr>
<td>13</td>
<td>Darlene’s Dolls</td>
<td>28</td>
<td>Sewer’s Center</td>
</tr>
<tr>
<td>14</td>
<td>Fleisch Realty</td>
<td>29</td>
<td>Tire Specialties</td>
</tr>
<tr>
<td>15</td>
<td>Hernandez Electronics</td>
<td>30</td>
<td>Von’s Video Store</td>
</tr>
</tbody>
</table>

**Step 2:** Start looking anywhere in Table A and record five two-digit groups. Ignore all two-digit groups greater than 30!

**Step 3:** Record the five clients represented by your SRS.
Chapter 2 Exercises

1. There are 20 students in my class. They sit in assigned seats, consisting of 4 rows of 5 students each. I want to take a simple random sample consisting of 4 of the students in my class. To do this, I select a single student from each row as follows. I write the numbers 1 to 5 on identical slips of paper. I mix the slips in a hat and draw one at random. I count this number of seats in from the left in the front row and select the student in this seat. For example, if the number selected is 3, I select the third student in from the left in the first row. I replace the slip in the hat, again mix the slips, and draw a new number. The student seated this many seats in from the left in the second row is selected. I repeat this process for the remaining two rows. Every student in the class has a 1-in-5 chance of being selected when I come to their row. Thus, every student has the same chance of being selected. Is this sample an SRS?

No. Each group of 5 students is not equally likely - it’s not possible to choose four students from the same row.

2. Design your own bad sample. Your college wants to gather student opinion about parking for students on campus. It isn’t practical to contact all students.

(a) Give an example of a way to choose a sample of students that is poor practice because it depends on voluntary response.

Send an email to students to ask about parking preferences.

(b) Give another example of a bad way to choose a sample that doesn’t use voluntary response.

Interviewing students as they enter the parking garage.