Event

• An event is something that either happens or doesn't happen, or something that either is true or is not true

• Examples:
  • You get cancer
  • A randomly selected person is obese
  • A particular mutation occurs
  • It snows tonight

Probability

• The probability of event A, P(A), is the probability that A will happen
• Probability always refers to an event
• Probability is always between 0 and 1
  • P(A) = 1 means A will definitely happen
  • P(A) = 0 means A will definitely not happen

Questions of the Day

How common is breast cancer?

How does breast cancer risk depend on other variables?

Ways of Expressing Probability

• 1 in 8 women will get breast cancer
• 1/8 of women will get breast cancer
  • The proportion of women who will get breast cancer is 1/8, or 0.125.
• 12.5% of women will get breast cancer
• The probability of breast cancer for a female is 0.125
  • These statements are all saying the same thing
### Relative Frequency Table

- Probability statements can be directly translated into a relative frequency table:

<table>
<thead>
<tr>
<th>Get Breast Cancer</th>
<th>Do not Get Breast Cancer</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125</td>
<td>0.875</td>
<td>1</td>
</tr>
</tbody>
</table>

### Frequency Table

- Probability statements can also be translated into a frequency table, although unlike data description, the total is arbitrary:

<table>
<thead>
<tr>
<th>Get Breast Cancer</th>
<th>Do not Get Breast Cancer</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Get Breast Cancer</th>
<th>Do not Get Breast Cancer</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>125</td>
<td>875</td>
<td>1000</td>
</tr>
</tbody>
</table>

- Either of these are equally valid for probability calculations!

### Odds

If \( p \) denotes the probability of an event, the **odds** are defined as

\[
\text{odds} = \frac{p}{1-p}
\]

- Interpreting odds
  - Odds of 1 indicate 50/50
  - \( p < 0.5 \) yield odds < 1
  - \( p > 0.5 \) yield odds > 1
- Odds of 3, or 3:1, mean that out of 4 times, we would expect the variable to be in that category 3 times and out of that category 1 time

### Breast Cancer Odds

\[
\text{odds} = \frac{1/8}{1-1/8} = \frac{1/8}{7/8} = \frac{1}{7}
\]

- Commonly expressed as 1:7 or 1/7
- The odds of a woman getting breast cancer are 1:7
- For every one woman who will get breast cancer, 7 women will not.

### Conditional Probability

- \( P(A \mid B) \) is the probability of \( A \) if we know \( B \) has happened or is true
- This is read in multiple ways:
  - “probability of \( A \) if \( B \)”
  - “probability of \( A \) given \( B \)”
  - “probability of \( A \) conditional on \( B \)”
- You may also see this written as \( P(A \mid B) \)

### Conditional Probability

- The probability and odds that we calculated are restricted only to females, so we are implicitly **conditioning** on the fact that gender is female.
- **For females**, what’s the probability of getting breast cancer?
- What proportion of women will get breast cancer?
- **Conditional probability** is the probability of an event, conditional on (or given) that another variable takes a specific value (gender = female)
• P(survival if advanced stage) = 0.27
• P(survival if early detection) = 0.98

What does 0.43% represent?

a) P(breast cancer if 30 – 39 years old)

b) P(30 – 39 years old if breast cancer)

What does this tell us?

a) P(breast cancer if first-degree relative)

b) P(first degree relative if breast cancer)

What does this tell us?

a) P(breast cancer if first-degree relative)

b) P(first degree relative if breast cancer)

c) P(breast cancer if no family history) = 0.15

These are given as facts, but clearly there is some uncertainty! (or different populations being studied...)
• P(breast health routine if female) = 0.599
• P(breast health routine if male) = 0.212

---

**Conditional Probability**

• P(breast cancer if female) = 0.125
• What’s P(female if breast cancer)? 0.99
• P(A if B) is NOT the same as P(B if A)!!!
• This is an important point, and something that is easily confused. Be careful with conditioning!

---

**Odds Ratio**

The odds ratio (OR) is the ratio of the odds in one group to the odds in the other group:

\[
OR = \frac{p_1(1-p_2)}{p_2(1-p_1)}
\]

• Odds ratios of 1 indicate no difference between the groups (no relationship between the two variables)

---

**Odds Ratio**

\[
OR = \frac{odds of getting breast cancer for males}{odds of getting breast cancer for females}
\]

\[
P(\text{breast cancer if female})
\]

\[
\frac{1 - P(\text{breast cancer if female})}{P(\text{breast cancer if male})}
\]

\[
\frac{1}{1 - \frac{1}{8}} = \frac{7}{799} = 114.14
\]

\[
\frac{1 - 1/800}{1/800}
\]

---

**Cats and Schizophrenia**

Table 1: Cat ownership in NAMI families and controls.

<table>
<thead>
<tr>
<th>Case</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>80/185</td>
<td>89/199</td>
</tr>
<tr>
<td>69/252</td>
<td>70/255</td>
</tr>
<tr>
<td>12/120</td>
<td>13/132</td>
</tr>
</tbody>
</table>

-produces are derived from chis squares, 2-tailed, ORs shown as mean (95% CI)
Breast Cancer Screening

- 1% of women at age 40 who participate in routine screening have breast cancer.
- 80% of women with breast cancer get positive mammographies.
- 9.6% of women without breast cancer get positive mammographies.
- A 40-year old woman participates in routine screening and has a positive mammography. What's the probability she has cancer?

What is this asking for?

a) \( P(\text{cancer} \mid \text{positive mammography}) \)

b) \( P(\text{positive mammography} \mid \text{cancer}) \)

c) \( P(\text{positive mammography} \mid \text{no cancer}) \)

d) \( P(\text{positive mammography}) \)

e) \( P(\text{cancer}) \)

---

100,000 women in the population

- 1% have cancer
- 99% are cancer-free

800 test positive
200 test negative
9,504 test positive
89,496 test negative

Thus, \( \frac{800}{800 + 9,504} = 7.8\% \) of positive results have cancer (you could also get this by creating a two-way table)
To Do

- Do HW 2.1 (due Friday, 9/25)