"Million-Dollar Murray" pg. 64-73 in reader

弱势群体的健康支出

N = 1,000

平均支出: $1,000

如果直方图看起来像这样:

Then Reno could cut the total in 1/2 by changing policies so that the mean went down from $1,000 to $500.

例如，通过获取大量这些人的数据，但直方图实际上看起来像这样:

Completely diff. policy cuts the bill almost to $0: find Murray a job (!).

Sampling Review:

1) Why did the large sample size of LD not ensure a correct answer? A: w/ un-biased data-gathering plan, more data = more accuracy, but w/ a biased sampling plan, not true any more b/c larger sample just perpetuates the bias.

2) Ex. Delaware (500,000 pop.) & New York (15,000,000 pop.)

\[ N = \text{pop. size} = N_D = N_{NY} \]

Suppose Gallup poll: take random sample of \( N_D = 1,000 \) & \( N_{NY} = 1,000 \).

Accuracy of \( \text{Del} \) poll > accuracy of \( \text{NY} \) poll \( \frac{1}{3} \) right answer not many
Actual fact: accuracy of random sample determined by \( \sqrt{n} \), not by \( N \) or \( N/N \) (as long as \( n \) is a lot smaller than \( N \) (\( n \ll N \)) & the sampling is SRS (at random w/ out replacement).
IV. Probability

C.S. #7: Tay-Sachs Disease (pg. 74)

\[
P(\text{rain tomorrow}) = 30\% = \frac{\text{# of days it actually rained}}{\text{total # of days in past "like" tomorrow w/ respect to relevant variables for rain}} \leq \frac{1000}{1000} = 100\%
\]

Ex) equally likely model (ELM)

\[
P(\text{H in toss of a fair coin}) = \frac{1}{2} = 50\%
\]

\[
\begin{array}{c|c|c}
\text{random} & \text{sample} & n=1 \\
\hline
1 & ELM & [Y_1] \\
2 & [Y_1] & n=1 \\
\end{array}
\]

\[
p(Y_1 = 9) = \frac{1}{3} = 33\%
\]

\[
p(Y_1 = \text{odd#}) = \frac{2}{3} \approx 67\%
\]

STARTS HERE

<table>
<thead>
<tr>
<th>father</th>
<th>mother</th>
<th>offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>D</td>
<td>r</td>
<td>D</td>
</tr>
<tr>
<td>r</td>
<td>D</td>
<td>r</td>
</tr>
</tbody>
</table>

\[\text{note:}\]

\[D=\text{dominant}\]

\[r=\text{recessive}\]

\[\leftarrow 2 \text{ phenotypes can be observed}\]

Ex) snapdragon color:

\[\begin{array}{c|c|c}
\text{f.} & \text{m.} & \text{offspring} \\
\hline
\text{white} & \text{white} & \text{white} \\
\text{white} & \text{red} & \text{pink} \\
\text{red} & \text{white} & \text{pink} \\
\text{red} & \text{red} & \text{red} \\
\end{array}\]

\[\text{Carriers (50%)}\]

\[\text{Tay-Sachs (0%)}\]

\[\text{normal (carrier)}\]

\[\text{3 phenocalled "additive model"}\]

\[\text{if F,M are both carriers}\]

\[\text{genetic name}\]

\[\text{A,A}\]

\[\text{informal name}\]

\[\text{normal or non-carrier}\]

\[100\% \text{ enzyme change}\]
26/4 AMS 5

<table>
<thead>
<tr>
<th>genetic name</th>
<th>genes</th>
<th>informal name</th>
<th>enzyme</th>
<th>chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>heterozygous</td>
<td>A,a</td>
<td>carriers</td>
<td>50%</td>
<td>( \frac{3}{4} = 50% )</td>
</tr>
<tr>
<td>homozygous</td>
<td>aa</td>
<td>T-S</td>
<td>0%</td>
<td>( \frac{1}{4} = 25% )</td>
</tr>
</tbody>
</table>

\[ P(1 \text{ or more T-S in 5 kids}) = ? \]

if ELM then \( \frac{5}{6} = 83\% \) but ELM doesn't apply

but ELM doesn't apply because expected \( \# \) of T-S kids is around \( 1 \left( \frac{5}{4} \right) \to 1, 2.5 \)

\[ \text{NOT around 2.5} \]

\[ \text{# of T-S kids} \]

\[ \begin{array}{c}
0 \\
1 \\
2 \\
3 \\
4 \\
5 \\
\end{array} \]

\[ P(1 \text{ or more T-S}) = P(1 \text{ T-S or... 2 or... or 5}) \]

↑ learn next time!