In [1]:
```
%%html
<style>
    .gray {
        background-color: #dfe0e8;
    }
</style>
```

Probability Practice

Part A. Visitors to your website are asked to answer a single survey question before they get access to the content on the page. Among all of the users, there are two categories: Random Clicker (RC), and Truthful Clicker (TC). There are two possible answers to the survey: yes and no. Random clickers would click either one with equal probability. You are also giving the information that the expected fraction of random clickers is 0.3. After a trial period, you get the following survey results: 65\% said Yes and 35\% said No. What fraction of people who are truthful clickers answered yes? Hint: use the rule of total probability.

Solving for $P(CY|TC)$.

Rule of total probability:

$$P(A) = \sum_{i=1}^{N} A \cap B_i = \sum_{i=1}^{N} P(B_i)P(A|B_i)$$

$$P(B_i, B_j) = 0 \;\; (i \neq j)$$

$$\sum_{i=1}^{N} P(B_i) = 1$$

Note: I'm using CY as shorthand for click yes and CN as shorthand for click no Given:

- $P(CY|RC) = 0.5$ and $P(CN|RC) = 0.5$
- $P(RC) = 0.3$ and $P(TC) = 0.7$
- $P(CY) = 0.65$ and $P(CN) = .35$

$$P(CY|TC) = \frac{P(CY \cap TC)}{P(TC)}$$

Need to determine $P(CY \cap TC)$ by using total probability theorem.

$$P(CY) = P(CY \cap RC) + P(CY \cap TC)$$

$$P(CY \cap TC) = P(CY) - P(CY \cap RC) = P(CY) - P(RC)P(CY|RC) =$$

$$0.65 - 0.3(0.5) = 0.5$$

$$P(CY|TC) = \frac{P(CY \cap TC)}{P(TC)} = \frac{0.5}{0.7} = 71.4$$

71.4% of the people who are truthful clickers click yes.

Part B. Imagine a medical test for a disease with the following two attributes:

- The sensitivity is about 0.993. That is, if someone has the disease, there is a probability of 0.993 that they will test positive.
- The specificity is about 0.9999. This means that if someone doesn't have the disease, there is probability of 0.9999 that they will test negative.
- In the general population, incidence of the disease is reasonably rare: about 0.0025% of all people have it (or 0.000025 as a decimal probability).

Suppose someone tests positive. What is the probability that they have the disease?

Note:

- TP = Test postivie
- TN = Test negative
- D = disease
- D^c = no disease

The givens are:

- $P(TP|D) = 0.993$
- $P(TN|D^c) = 0.9999$
- $P(D) = 0.000025$
- $P(D^c) = 1 - P(D) = 0.999975$

Find P(D|TP)

I know:

$$P(D|TP) = \frac{P(D \cap TP)}{P(TP)} = \frac{P(TP|D)\,P(D)}{P(TP)} = \frac{0.993\,(0.000025)}{?}$$

I have to solve for P(TP) before I can finish filling in this equation.

$$P(TP) = P(TP \cap D) + P(TP \cap D^c) = P(TP|D)\,P(D) + P(TP|D^c)\,P(D^c)$$

I have everything except for $P(TP|D^c)$ in the given information, but this can easily be derived by: $P(TP|D^c) = 1 - P(TN|D^c) = 1 - 0.9999 = 0.0001$

Therefore:

$$P(TP) = P(TP \cap D) + P(TP \cap D^c) =$$

$$0.993\,(0.000025) + 0.0001\,(0.999975) = 0.0001248225$$

Finally,

$$P(D|TP) = \frac{P(D \cap TP)}{P(TP)} = \frac{P(TP|D)\,P(D)}{P(TP)} =$$

$$\frac{0.993\,(0.000025)}{0.0001248225} = 0.19888$$

Given that someone tests positive, they have a 19.89% of actually having the disease.

In [ ]: