Example (poll). Select n (yes/no) voters at random from a population of size N. Let $p = M/N$ be the proportion of "yes" voters. Estimate p by $\hat{p} = X/n$ with X = number of yes responses in sample.

We consider sampling with replacement (SWR): there are N^n equally likely possible samples (ordered). Number of samples with x "yes" ($n - x$ "no"):

$$M^x \times (N-M)^{n-x} \times \#\,\text{ways to place } x \text{ "yes" among } n \text{ responses.}$$

$$\begin{aligned} P(X = x) &= \frac{\binom{n}{x} M^x (N-M)^{n-x}}{N^n} = \binom{n}{x}\left(\frac{M}{N}\right)^x \left(\frac{N-M}{N}\right)^{n-x} \\ &= \binom{n}{x} p^x (1-p)^{n-x}. \end{aligned}$$

Chance of estimating p within a small value Δ:

$$\begin{aligned} P(p - \Delta \le \hat{p} \le p + \Delta) &= P\{n(p-\Delta) \le X \le n(p+\Delta)\} \\ &= \sum_{n(p-\Delta) \le x \le n(p+\Delta)} \binom{n}{x} p^x (1-p)^{n-x}. \end{aligned}$$

1.4 Conditional probability and Bayes' Theorem

Example. Suppose 40% of a population has a hidden disease D. A test is available which gives 5% false positives and 20% false negatives. What does this mean?

Let N be the population size and P the event "positive response". Then $0.4N$ "people" have D and 20% of 40% (or of $0.4N$) will fail to show P, i.e. $0.2 \cdot 0.4N = 0.08N$. Therefore $P(D \text{ and not } P) = P(DP^c) = \frac{0.08N}{N} = 0.08$, which is relative to the entire population $[= P(P^c|D)P(D)]$, whereas the 20% is relative to the diseased subpopulation $[P(P^c|D) = P(P^cD)/P(D)]$. It can be regarded as a conditional probability.

Definition. Let A, B be events. The **conditional probability** of A given B is

$$P(A|B) = \frac{P(AB)}{P(B)} \quad \text{if } P(B) > 0.$$

Example (dice). $P(\text{"1"}|\text{odd}) = \dfrac{1/6}{1/2} = \dfrac{1}{3} = \dfrac{\#\{\text{"result 1 and odd"}\}}{\#\{\text{"odd"}\}}$ with "odd" $= \{\mathbf{1}, 3, 5\}$.

Example (hidden disease). $20\% = P(P^c|D)$ (false negative), $5\% = P(P|D^c) = P(PD^c)/P(D^c)$ (false positive), so $P(PD^c) = P(P|D^c)P(D^c) = 0.05 \cdot 0.6 = 0.03$.

Theorem (multiplication rule).
(i) $P(AB) = P(B)P(A|B)$,
(ii) $P(A_1A_2 \ldots A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \ldots A_{n-1})$.

Example (cont.). Suppose that the false positive/negative rates are known (5% and 20%) and that 35% of the population tests positive. What is the (unknown) disease rate $P(D)$?

By assumption we know $P(P) = 0.35, P(P|D^c) = 0.05$ and $P(P^c|D) = 0.2$. Therefore $0.35 = P(P) = P(PD) + P(PD^c) = P(P|D)P(D) + P(P|D^c)P(D^c) = \{1 - P(P^c|D)\}P(D) + P(P|D^c)\{1 - P(D)\} = (1 - 0.2)P(D) + 0.05\{1 - P(D)\}$, i.e., $0.35 = 0.8P(D) + 0.05 - 0.05P(D)$ so $P(D) = 0.4$.

Corollary.
(i) $P(B|B) = 1$,
(ii) $A \subset B \Rightarrow P(A|B) = P(A)/P(B)$,
(iii) $AB = \emptyset \Rightarrow P(A|B) = 0$.

Theorem. Suppose $B_1, B_2, \ldots$ is a partition of $\mathcal{S}$.
(i) $P(A) = \sum_n P(A|B_n)P(B_n)$ (**total probability**),
(ii) $P(B_i|A) = \dfrac{P(A|B_i)P(B_i)}{\sum_n P(A|B_n)P(B_n)}$ (**Bayes' Theorem**).

Proof. (i) Use $P(A) = \sum_n P(AB_n)$ and the multiplication rule.
(ii) Use $P(B_i|A) = P(B_iA)/P(A)$ and then (i) for the denominator and the multiplication rule for the numerator. □

Example. Suppose 60% of marriages have children, 40% end in divorce, and 20% of marriages with children end in divorce. Do $0.4 \cdot 0.6 = 0.24 = 24\%$ of marriages end in divorce with children?
Answer: No! $P(DC) = 0.24$ must be wrong since $P(DC) = P(D|C)P(C) \leq P(D|C) = 0.2$. Correct value: $P(DC) = P(D|C)P(C) = 0.2 \cdot 0.6 = 0.12$

How many marriages (%) without children divorce?
Answer: $P(D|C^c) = P(DC^c)/P(C^c) = \{P(D) - P(DC)\}/\{1 - P(C)\} = (0.4 - 0.12)/(1 - 0.6) = 0.7 = 70\%$.

Theorem. Suppose $P(B) > 0$ and define $P^B(A) = P(A|B)$. Then P^B is a (valid) probability measure.

Proof. (i) $P^B(A) \geq 0$ by definition.
(ii) $P^B(\mathcal{S}) = P(B\mathcal{S})/P(B) = P(B)/P(B) = 1$.
(iii) Let $A_1, A_2, \ldots$ be disjoint. Use that P is a probability measure and therefore countably additive to obtain $P^B(\cup A_n) = P(\cup A_n B)/P(B) = \sum P(A_n B)/P(B) = \sum P^B(A_n)$. □

Conclusion: Calculate with P^B "as usual", e.g. $P(A^c|B) = P^B(A^c) = 1 - P^B(A) = 1 - P(A|B)$.

Example. Roll a dice until a 6 appears. Given that a 6 appears on roll $n+1$, what is the distribution for the observed number of 1's? *Solution:* consider "new" (conditional) sample space: n selections with replacement from $\{1, 2, 3, 4, 5\}$. Compare example "poll" ($p = M/N$) with here $p = 1/5$:

$P($"1 appears x times"$|$ "6 appears first on the $(n+1)^{st}$ roll"$) = \binom{n}{x}\left(\frac{1}{5}\right)^x\left(\frac{4}{5}\right)^{n-x}$, $x \in \{0, 1, \ldots, n\}$.

Example. Two colleagues have learned: if either forgets her umbrella then the chance of rain is 2/3, if neither does it is 1/3. Suppose that each forgets her umbrella 1/2 of the time and both simultaneously 1/4 of the time. What is the chance of rain if both forget their umbrellas?
Solution. Let $A_i =$ "i of them forgot umbrella", $i = 0, 1, 2$, $R =$ "rain". What is $P(R|A_2)$?
We have $P(A_2) = 1/4, P(A_1) = 1/2, P(A_0) = 1/4$ and $P(R|A_0) = 1/3, P(R|A_1 \cup A_2) = 2/3$. Then $P(R|A_2) = P(RA_2)/P(A_2) = P(RA_2) \times 4$ with $P(RA_2) =?$. $P(RA_2)$ is restricted, $P(RA_2) \leq P(A_2) = 1/4$, so $P(R|A_2) = P(RA_2) \times 4 \leq 1/4 \cdot 4 = 1$. It can be anything between 0 and 1.