# Section 1: Introduction to probability (solutions)

ARE 210

August 29, 2017

3.1) **Conditional independence:** Define $A_1$ and $A_2$ to be independent conditional on the event $B$ occurring if $P(A_1 \cap A_2|B) = P(A_1|B)P(A_2|B)$. Does independence conditional on each event in a partition of the sample space imply independence? If so, why? If not, construct a counterexample.

Let $\{B_i\}_{i=1}^N$ partition the sample space $\Omega$.

$$\begin{aligned}
P(A_1 \cap A_2) &= \sum_{i=1}^{N} P(A_1 \cap A_2 \cap B_i) \\
&= \sum_{i=1}^{N} P(A_1 \cap A_2|B_i)P(B_i) \\
&= \sum_{i=1}^{N} P(A_1|B_i)P(A_2|B_i)P(B_i)
\end{aligned}$$

We can see that $P(A_i|B_i) = P(A_i)$ would be sufficient for the result to hold. Can we construct an example where this doesn't hold?

Let $\Omega = \{\text{fair coin chosen}, \text{two headed coin chosen}\} \times \{0,1\}^2$ be the sample space, representing the choice of a coin (fair or two headed), and the possible outcomes of two coin flips. Let $P(\text{fair coin chosen}) = 0.5$, $P(\{1,1\}|\text{two headed coin chosen}) = 1$, and all possible sequences of coin flips have equal probability if the fair coin is chosen. Let $A_1 = \text{heads first flip}$ and $A_2 = \text{heads second flip}$. In this case

$$\begin{aligned}
P(A_1|B_1)P(A_2|B_1) &= 0.25 \qquad P(A_1 \cap A_2|B_1) = 0.25 \\
P(A_1|B_2)P(A_2|B_2) &= 1 \qquad P(A_1 \cap A_2|B_2) = 1
\end{aligned}$$

we have conditional independence on each event in a partition of the sample space ($\{B_1, B_2\}$ partitions $\Omega$), but

$$P(A_1)P(A_2) = \frac{9}{16} \qquad P(A_1 \cap A_2) = \frac{5}{8}$$

we do not have independence.

3.2) **Practice constructing counterexamples:** Recall our formula for Bayesian updating, $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$. Suppose we observe the events $B_1, B_2, \ldots, B_N$. Consider calculating $P(A|\cap_{i=1}^N B_i)$ via **batch Bayesian updating** and **sequential Bayesian updating**. Define the batch Bayesian updating approach by

$$P(A|\cap_{i=1}^N B_i) = \frac{P(\cap_{i=1}^N B_i|A)P(A)}{P(\cap_{i=1}^N B_i)}$$

Define the sequential Bayesian updating approach recursively by

$$P^{SEQ}(A|B_1) = \frac{P(B_1|A)P(A)}{P(B_1)}$$
$$P^{SEQ}(A|\cap_{i=1}^K B_i) = \frac{P(B_K|A)P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)}{P(B_K)}$$

a) Propose two sufficient conditions for the two approaches to yield the same answer. Which approach do you think is "correct"? Hint: $\frac{P(B_K|A)P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)}{P(B_K)} = \left(\frac{P(B_K|A)}{P(B_K)}\right) P^{SEQ}(A|\cap_{i=1}^{K-1} B_i)$.

We can rewrite $P^{SEQ}(A|\cap_{i=1}^N B_i) = \frac{\prod_{i=1}^N P(B_i|A)}{\prod_{i=1}^N P(B_i)} P(A)$. Then we can see that 1) $B_i$ are independent conditional on $A$, and 2) $B_i$ are independent are sufficient conditions for $P^{SEQ}(A|\cap_{i=1}^N B_i) = P(A|\cap_{i=1}^N B_i)$.

The batch approach returns the correct result in that it strictly applies Bayes rule. The sequential approach, as we've formulated it, in each recursive step throws out the information $\int_{i=1}^{K-1} B_i$ gives us about $P(B_K|A)$ and $P(B_K)$.

b) How can you correct the sequential (batch) approach so it returns the same answer as the batch (sequential) approach?

To correct the sequential approach, we need to avoid throwing out that information. Redefine the sequential approach by

$$P^{SEQ'}(A|B_1) = P^{SEQ}(A|B_1)$$
$$P^{SEQ'}(A|\cap_{i=1}^K B_i) = \frac{P(B_K|A \cap (\cap_{i=1}^{K-1} B_i))P^{SEQ'}(A|\cap_{i=1}^{K-1} B_i)}{P(B_K|\cap_{i=1}^{K-1} B_i)}$$

We can now show by induction that the two approaches are identical. First, $P(A|B_1) = P^{SEQ}(A|B_1) = P^{SEQ'}(A|B_1$. Second, suppose $P^{SEQ'}(A|\cap_{i=1}^{K-1} B_i) = P(A|\cap_{i=1}^{K-1} B_i)$ for some $K$. Then,

$$\begin{aligned}
P^{SEQ'}(A|\cap_{i=1}^{K} B_i) &= \frac{P(B_K|A\cap(\cap_{i=1}^{K-1}B_i))P(A|\cap_{i=1}^{K-1} B_i)}{P(B_K|\cap_{i=1}^{K-1} B_i)} \\
&= \frac{\frac{P(B_K\cap A\cap(\cap_{i=1}^{K-1}B_i))}{P(A\cap(\cap_{i=1}^{K-1}B_i))}\frac{P(A\cap(\cap_{i=1}^{K-1}B_i))}{P((\cap_{i=1}^{K-1}B_i))}}{\frac{P(B_K\cap(\cap_{i=1}^{K-1}B_i))}{P((\cap_{i=1}^{K-1}B_i))}} \\
&= \frac{P(B_K\cap A\cap(\cap_{i=1}^{K-1}B_i))}{P(B_K\cap(\cap_{i=1}^{K-1}B_i))} \\
&= P(A|\cap_{i=1}^{K} B_i)
\end{aligned}$$

c) For $N = 2$, construct an example where the two approaches yield different answers, but the second condition is not violated.

Use the same counterexample and sample space as in 3.1), letting $B_1$ = heads first flip and $B_2$ = heads second flip, $A$ = fair coin chosen. We have shown than $B_1$ and $B_2$ are independent conditional on $A$, but they are not independent.

Batch updating, $P(A|B_1 \cap B_2) = \frac{P(B_1\cap B_2|A)P(A)}{P(B_1\cap B_2)} = \frac{\frac{1}{4}\frac{1}{2}}{\frac{5}{8}} = \frac{1}{5}$.

Sequential updating, $P^{SEQ}(A|B_1) = \frac{P(B_1|A)P(A)}{P(B_1)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{3}{4}} = \frac{1}{3}$. $P^{SEQ}(A|B_1 \cap B_2) = \frac{P(B_2|A)P^{SEQ}(A|B_1)}{P(B_2)} = \frac{\frac{1}{2}\frac{1}{3}}{\frac{3}{4}} = \frac{2}{9}$.

Modified sequential updating, $P^{SEQ'}(A|B_1) = \frac{1}{3}$. $P^{SEQ'}(A|B_1\cap B_2) = \frac{P(B_2|A\cap B_1)P^{SEQ'}(A|B_1)}{P(B_2|B_1)} = \frac{\frac{1}{2}\frac{1}{3}}{\frac{5}{6}} = \frac{1}{5}$.

d) For $N = 2$, construct an example where the two approaches yield different answers, but the first condition is not violated.

Let $B_1$ = heads first flip and $B_2$ = heads second flip. Suppose two coins can be chosen, a sticky coin ($P(B_1|A) = \frac{1}{2}$, $P(B_2|B_1 \cap A) = 1$, and $P(B_2^c|B_1^c \cap A) = 1$, where $A$ = sticky coin chosen) and a flippy coin ($P(B_1|A^c) = \frac{1}{2}$, $P(B_2|B_1 \cap A^c) = 0$ and $P(B_2^c|B_1^c \cap A^c) = 0$). Let $\Omega = \{\text{sticky coin chosen}, \text{flippy coin chosen}\} \times \{0, 1\}^2$, and assume $P(A) = \frac{1}{2}$. $B_1$ and $B_2$ are not independent conditional on $A$, but one can show $B_1$ and $B_2$ are independent.

Batch updating, $P(A|B_1 \cap B_2) = \frac{P(B_1\cap B_2|A)P(A)}{P(B_1\cap B_2)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{1}{4}} = 1$.

Sequential updating, $P^{SEQ}(A|B_1) = \frac{P(B_1|A)P(A)}{P(B_1)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}$. $P^{SEQ}(A|B_1 \cap B_2) =$

$\frac{P(B_2|A)P^{SEQ}(A|B_1)}{P(B_2)} = \frac{\frac{1}{2}\frac{1}{2}}{\frac{1}{2}} = \frac{1}{2}$.

Modified sequential updating, $P^{SEQ'}(A|B_1) = \frac{1}{2}$. $P^{SEQ'}(A|B_1 \cap B_2) = \frac{P(B_2|A,B_1)P^{SEQ'}(A|B_1)}{P(B_2|B_1)} = \frac{1\frac{1}{2}}{\frac{1}{2}} = 1$.