Probabilities and statistics
Lecturer: Prof. Dr. Sara van de Geer
Prof. Dr. Martin Larsson

Serie 4

March 16th, 2015

Q1. Define $B_0 = \frac{1}{2}$. We want to play the following game: "In time n we bet $B_n = 2B_{n-1}$. Then we flip a coin X_n if it's -1 (Tail) we lose all the money we bet, if it's 1 (Head) we win the same amount of money we bet and we stop the game". Define $(V_n)_{n\in\mathbb{N}}$ the amount of money bet at time n:

(a) Compute the distribution of profit $(VX)_n$. What is the probability of losing?.

(b) Compute the expected value of $(VX)_n$. Then compute

$$\mathrm{Var}((VX)_n) := \mathbb{E}\left(((VX)_n - (\mathbb{E}(VX)_n))^2\right).$$

Solution

(a) Define T the time where we stop the game, i.e., $T = \inf\{n \in \mathbb{N} : X_n = 1\}$.We have that:

$$\begin{aligned}
(V \cdot S)_n &= \sum_{k=1}^{n} V_k X_k = \sum_{k=1}^{n} 2^{k-1} 1_{\{T\geq k\}} X_k \\
&= \left(\sum_{k=1}^{T} 2^{k-1} X_k\right) 1_{\{T<\infty\}} - \left(\sum_{k=1}^{n} 2^{k-1}\right) 1_{\{T=\infty\}} \\
&= \left(\sum_{k=1}^{T-1} 2^{k-1}(-1) + 2^{T-1}\right) 1_{\{T<\infty\}} + (1-2^n) 1_{\{T=\infty\}} \\
&= (1 - 2^{T-1} + 2^{T-1}) 1_{\{T<\infty\}} + (1-2^n) 1_{\{T=\infty\}} \\
&= 1_{\{T<\infty\}} + (1-2^n) 1_{\{T=\infty\}},
\end{aligned}$$

thus $(V \cdot S)_n$ can take only 2 values: 1 and $1-2^n$. Then for computing the probabilities we have to do the following:

$$\mathbb{P}[(V\cdot S)_n = 1 - 2^n] = \mathbb{P}[T = \infty] = \mathbb{P}\left[\bigcap_{i=1}^{n}\{X_i = -1\}\right] = 2^{-n},$$

$$\mathbb{P}[(V\cdot S)_n = 1] = \mathbb{P}[T < \infty] = 1 - \mathbb{P}[T = \infty] = 1 - 2^{-n}.$$

Finally, the probability of losing is 2^{-n}.

(b) Computing the expected value:

$$\mathbb{E}[(V\cdot S)_n] = (1-2^n)2^{-n} + 1(1-2^{-n}) = 0.$$

To compute the variance:

$$\begin{aligned}
Var[(V\cdot S)_n] &= \mathbb{E}[(V\cdot S)_n^2] - \mathbb{E}[(V\cdot S)_n]^2 \\
&= \mathbb{E}[(V\cdot S)_n^2] \\
&= (1-2^n)^2 2^{-n} + 1^2(1-2^{-n}) \\
&= 2^{-n} - 2 + 2^n + 1 - 2^{-n} = 2^n - 1.
\end{aligned}$$

Q2. We are interested in studying the probability of success of a student at an entrance exam to two departments of a university. Consider the following events

$$\begin{aligned} A &= \{\text{The student is man}\}, \\ A^c &= \{\text{The student is woman}\}, \\ B &= \{\text{The student applied for department I}\}, \\ B^c &= \{\text{The student applied for department II}\}, \\ C &= \{\text{The student was accepted}\}, \\ C^c &= \{\text{The student wasn't accepted}\}. \end{aligned}$$

We assume that we have the following probabilities (Berkeley 1973):

$\mathbb{P}(A) = 0.73$,
$\mathbb{P}(B \mid A) = 0.69$, $\mathbb{P}(B \mid A^c) = 0.24$,
$\mathbb{P}(C \mid A \cap B) = 0.62$, $\mathbb{P}(C \mid A^c \cap B) = 0.82$, $\mathbb{P}(C \mid A \cap B^c) = 0.06$, $\mathbb{P}(C \mid A^c \cap B^c) = 0.07$.

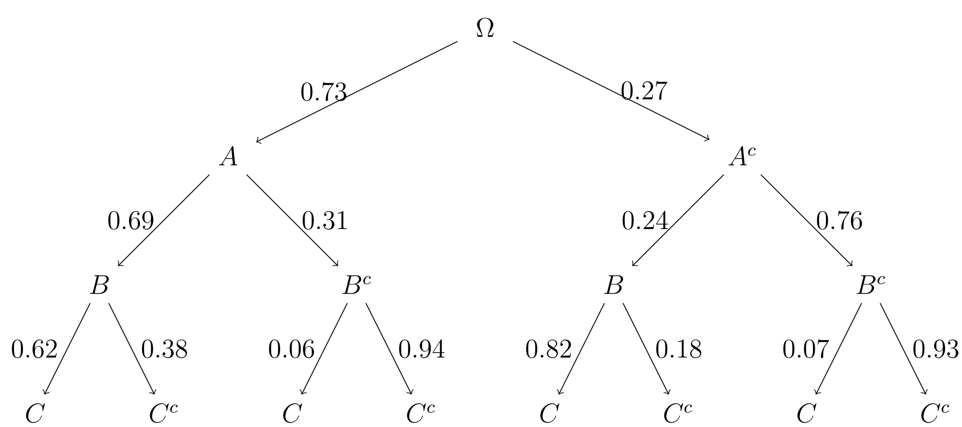
(a) Draw a tree describing the situation with the probabilities associated.

(b) Taking in to consideration the following probabilities $\mathbb{P}[C|A \cap B] = 0.62$, $\mathbb{P}[C|A^c \cap B] = 0.82$, $\mathbb{P}[C|A \cap B^c] = 0.06$, $\mathbb{P}[C|A^c \cap B^c] = 0.07$. With this information, do you think that in this examination women are disadvantaged?.

(c) Compute $\mathbb{P}(C \mid A)$ and $\mathbb{P}(C \mid A^c)$. Does this coincide with your answer of b)?.

Solution

(a) The tree can be drawn as:



(b) The probability of being accepted, given than you are a woman who postulated at the department I is $P(C \mid A^c \cap B) = 0.82$. That value is bigger than the probability of being accepted, given than you are a man who postulated at the department I, $\mathbb{P}(C \mid A \cap B) = 0.62$. This indicates that in department I females are not disadvantaged.

The probability of being accepted, given than you are a woman who postulated to the department is $P(C \mid A^c \cap B^c) = 0.07$. This value is bigger than the probability of being accepted given than you are a man who postulated at the department II $\mathbb{P}(C \mid A \cap B^c) = 0.06$. This indicates that in department II females are neither disadvantaged.

(c) Just computing

$$\mathbb{P}[C|A^c] = \frac{\mathbb{P}[C\cap A^c]}{\mathbb{P}[A^c]} = \frac{\mathbb{P}[C\cap A^c \cap B] + \mathbb{P}[C\cap A^c\cap B^c]}{\mathbb{P}[A^c]}$$
$$= \frac{0.82\cdot 0.27\cdot 0.24 + 0.07\cdot 0.27\cdot 0.76}{0.27} \sim 0.25,$$

and

$$\mathbb{P}[C|A] = \frac{\mathbb{P}[C\cap A]}{\mathbb{P}[A^c]} = \frac{\mathbb{P}[C\cap A \cap B] + \mathbb{P}[C\cap A\cap B^c]}{\mathbb{P}[A]}$$
$$= \frac{0.62\cdot 0.69\cdot 0.73 + 0.06\cdot 0.31\cdot 0.73}{0.73} \sim 0.45.$$

This shows that the percentage of women accepted are less than that of the men. This is not explained by the gender, but much more but the fact that women apply to the department with bigger rejection rate.

Q3. INTRODUCTION TO BAYESIAN STATISTICS We have m urns with red and white balls insede. The urn $i\in\{1,..m\}$ has $2i-1$ red balls and $2m-2i+1$ white ones. We randomly select an urn and extract with replacement n times. Define:

$$X_j := \begin{cases} 1 & \text{If the } j\text{-th ball is red,} \\ 0 & \text{If the } j\text{-th ball is white.}\end{cases}$$

We are interested in the following problem " Given that you see $(X_j)_{j=1}^n$, can you say from which urn the balls where taken?".

(a) Compute $\mathbb{P}(X_1 = x_1, ..., X_n = x_n)$ for $x_i \in \{0,1\}$. Are $X_1,..,X_n$ independent?.

(b) Compute the following probability:

$$\mathbb{P}\left(\text{The urn chosen is } i \mid X_1 = x_1, ..., X_n = x_n\right).$$

Show that this only depends on the number or red balls, i.e., $k = \sum_{i=1}^n x_i$.

(c) Compute $\mathbb{P}\left(\text{The urn chosen is } i \mid X_1 = x_1, ..., X_n = x_n\right)$ for $m = 3$ and $n = 3$:

	$i=1$	$i=2$	$i=3$
$k=0$			
$k=1$			
$k=2$			
$k=3$			

Solution

(a) Let $k = \sum_{j=1}^n x_j$ the amount of red balls taken out. Then

$$\mathbb{P}\left(X_1 = x_1, ..., X_n = x_n\right) = \sum_{i=1}^m \mathbb{P}\left(X_1 = x_1, ..., X_n = x_n \mid \text{Urn } i \text{ is chose}\right)\mathbb{P}\left(\text{Urn } i \text{ is chosen}\right)$$
$$= \sum_{i=1}^m \left(\frac{2i-1}{2m}\right)^k \left(\frac{2m-2i+1}{2m}\right)^{n-k} \frac{1}{m}.$$

We have that X_1 and X_2 are not independent because:

$$\begin{aligned}\mathbb{P}(\{X_1 = 1\} \cap \{X_2 = 1\}) &= \mathbb{P}(\{X_1 = 1\} \cap \{X_2 = 1\} \mid \text{Urn } i \text{ is chose})\, \mathbb{P}(\text{Urn } i \text{ is chose}) \\ &= \sum_{i=1}^{m} \left(\frac{2i-1}{2m}\right)^2 \frac{1}{2m} \\ &> \left(\sum_{i=1}^{m} \left(\frac{2i-1}{2m}\right) \frac{1}{m}\right)^2 = \mathbb{P}(X_1 = 1)\, \mathbb{P}(X_2 = 1).\end{aligned}$$

This happens because the first variable gives “information” about which urn we have chosen so it also gives information about the second variable.

(b) Just by definition:

$$\begin{aligned}&\mathbb{P}(\text{The urn chosen is } i \mid X_1 = x_1, ..., X_n = x_n) \\ =&\frac{\mathbb{P}(\text{The urn chosen is } i, X_1 = x_1, ..., X_n = x_n)}{\mathbb{P}(X_1 = x_1, ..., X_n = x_n)} \\ =&\frac{\frac{1}{m}\left(\frac{2i-1}{2m}\right)^k \left(\frac{2m-2i+1}{2m}\right)^k}{\sum_{j=1}^{m} \left(\frac{2j-1}{2m}\right)^k \left(\frac{2m-2j+1}{2m}\right)^k \frac{1}{m}} \\ =&\frac{\left(\frac{2i-1}{2m}\right)^k \left(\frac{2m-2i+1}{2m}\right)^k}{\sum_{j=1}^{m} \left(\frac{2j-1}{2m}\right)^k \left(\frac{2m-2j+1}{2m}\right)^k}.\end{aligned}$$

(c) We just have to compute and get:

	$i = 1$	$i = 2$	$i = 3$
$k = 0$	0.817	0.176	0.007
$k = 1$	0.439	0.474	0.088
$k = 2$	0.088	0.474	0.439
$k = 3$	0.007	0.176	0.817

Q4. MONTY HALL PROBLEM. Suppose you’re on a game show, and you’re given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what’s behind the doors, opens another door, say No. j ($j = 2, j = 3$), which has a goat. He then says to you, “Do you want to pick door No. l ($l = 2, l = 3$, $l \neq j$)?” Is it to your advantage to switch your choice?

For trying to solve this problem we define the following

$$\begin{aligned}B_i &= \text{“The car is behind the door i.”} && (i = 1, 2, 3), \\ A_j &= \text{“Moderator open door j.”} && (j = 2, 3).\end{aligned}$$

(a) Define a natural model for the problem. Under this model compute $\mathbb{P}(B_i)$ for $i \in \{1, 2, 3\}$ and $\mathbb{P}(A_j|B_i)$ for $i \in \{1, 2, 3\}$ and $j \in \{2, 3\}$.

(b) Compute, using the Bayes formula, $\mathbb{P}(B_1 \mid A_j)$ for $j \in \{2, 3\}$. Does opening the other door with a goat behind changes the probability that the car is behind door number 1?

(c) Would you change your choice?

Solution

(a) The model that represent the best is the problem is the one that gives the following probabilities:
$\mathbb{P}(B_1) = \mathbb{P}(B_2) = \mathbb{P}(B_3) = \frac{1}{3}$.
$\mathbb{P}(A_2|B_1) = \mathbb{P}(A_3|B_1) = \frac{1}{2}$.
$\mathbb{P}(A_2|B_2) = 0 \quad \mathbb{P}(A_3|B_2) = 1$.
$\mathbb{P}(A_2|B_3) = 1 \quad \mathbb{P}(A_3|B_3) = 0$.

(b) We just have to compute:

$$\begin{aligned}\mathbb{P}(B_1|A_2) &= \frac{\mathbb{P}(A_2|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A_2|B_1)\mathbb{P}(B_1) + \mathbb{P}(A_2|B_2)\mathbb{P}(B_2) + \mathbb{P}(A_2|B_3)\mathbb{P}(B_3)} \\ &= \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{2}\frac{1}{3} + 0\frac{1}{3} + 1\frac{1}{3}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.\end{aligned}$$

$$\begin{aligned}\mathbb{P}(B_1|A_3) &= \frac{\mathbb{P}(A_3|B_1)\mathbb{P}(B_1)}{\mathbb{P}(A_3|B_1)\mathbb{P}(B_1) + \mathbb{P}(A_3|B_2)\mathbb{P}(B_2) + \mathbb{P}(A_3|B_3)\mathbb{P}(B_3)} \\ &= \frac{\frac{1}{2}\frac{1}{3}}{\frac{1}{2}\frac{1}{3} + 1\frac{1}{3} + 0\frac{1}{3}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}.\end{aligned}$$

The probability of $\mathbb{P}(B_1 \mid A_2) = \mathbb{P}(B_1 \mid A_3) = \mathbb{P}(B_1)$ so it doesn't change.

(c) You have to change of door because $\mathbb{P}(B_1 \mid A_j) = \frac{1}{3}$, so $\mathbb{P}(B_l \mid A_j) = \frac{2}{3}$ where $1 \neq l \neq j$. This happens because you didn't get any additional information knowing that behind one of the doors that you didn't chose there was a goat.

Have a nice week ☺☻!!.