

Statistiek

Uitwerking inleveropgave 16 Oktober 2012

(Licht je antwoorden toe)

- 1 Jij en je bridgepartner krijgen beiden 13 kaarten uit een volledig pak van 52 kaarten.

- 1a Als jij precies 2 harten kaarten hebt, wat is de kans dat jouw partner geen harten heeft? A is de gebeurtenis dat jij precies twee harten krijgt, B dat jouw partner geen harten krijgt.

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{\binom{13}{2}\binom{39}{11}\binom{28}{13}}{\binom{52}{13}\binom{39}{13}}}{\frac{\binom{13}{2}\binom{39}{11}}{\binom{52}{13}}} = \frac{\binom{28}{13}}{\binom{39}{13}} = 0.004609727.$$

- 1b Als jij geen harten kaarten hebt, wat is de kans dat jouw partner precies 2 harten heeft? C is de gebeurtenis dat jouw partner precies twee harten krijgt, D dat jij geen harten krijgt.

$$P(C|D) = \frac{P(C \cap D)}{P(D)} = \frac{\frac{\binom{13}{2}\binom{39}{11}\binom{28}{13}}{\binom{52}{13}\binom{39}{13}}}{\frac{\binom{39}{13}}{\binom{52}{13}}} = \frac{\binom{13}{2}\binom{39}{11}\binom{28}{13}}{\binom{39}{13}\binom{39}{13}} = \frac{\binom{13}{2}\binom{26}{11}}{\binom{39}{13}} = 0.07419465.$$

- 2 Geef een voorbeeld van gebeurtenissen A en B zodat geldt

$$P(B) < P(A|B) < P(A).$$

Er wordt een dobbelsteen gegooid en de stochast X is het aantal ogen van de worp. A is de gebeurtenis “minder dan 6 gooien” en B de gebeurtenis “even gooien”. Dan geldt

$$\frac{1}{2} = P(B) < P(A|B) = \frac{2}{3} < P(A) = \frac{5}{6}.$$

- 3 Deze opgave bevat een fout: als er 10 balletjes zijn kan de verhouding rood/wit nooit 1 : 3 of 1 : 2 zijn. Gelukkig kan de opgave gewoon gemaakt worden door de informatie dat er 10 balletjes zijn te vergeten en te werken met de kansen $\frac{1}{4}$, $\frac{1}{3}$ en $\frac{2}{5}$, zoals de meeste van jullie deden. Sommige mensen vatten de $\frac{1}{3}$, $\frac{1}{2}$ en $\frac{2}{3}$ als de kans op een rood balletje op, dat heb ik niet fout gerekend.

Op een kermis staat een grote urn met rode en witte balletjes, 10 in totaal. Je mag tweemaal vijf balletjes pakken met terugleggen en op grond

daarvan raden wat de verhouding rode/witte balletjes in de urn is. Als je antwoord juist is ontvang je een prijs. Verder vertelt de eigenaar van de urn dat de verhouding rode/witte balletjes $\frac{1}{3}$, $\frac{1}{2}$ of $\frac{2}{3}$ is, maar hij zegt niet welk van de drie de juiste is. Aanvankelijk houdt je alledrie de hypothesen voor even waarschijnlijk en je besluit je antwoord te baseren op het tweemaal uitvoeren van een Bayesiaanse update.

De kansen op een rood balletje is dus $H_1 : p = \frac{1}{4}$, $H_2 : p = \frac{1}{3}$ of $H_3 : p = \frac{2}{5}$.

- 3a Je pakt 5 balletjes, waarvan er 2 rood zijn. Wat zijn de a-posteriori waarschijnlijkheden van de hypothesen op grond van deze informatie, en wat is de meest waarschijnlijke hypothese (de h_{MAP})?

Laat X het aantal rode balletjes na de eerste trekking zijn. Omdat de a-priori waarschijnlijkheden gelijk zijn geldt

$$P_1(H_i) = P_0(H_i | X = 2) = \frac{P_0(X = 2 | H_i)}{\sum_{j=1}^3 P_0(X = 2 | H_j)} = \frac{\binom{5}{2} p^2 (1-p)^3}{m},$$

waarbij $m = \binom{5}{2} \left(\frac{1}{4} \frac{3}{4} \frac{3}{4} + \frac{1}{3} \frac{2}{3} \frac{2}{3} + \frac{2}{5} \frac{3}{5} \frac{3}{5} \right)$. Dat geeft

$$P_1(H_1) = 0.2809533 \quad P_1(H_2) = 0.3507955 \quad P_1(H_3) = 0.3682511.$$

Dus $h_{\text{MAP}} = H_3$.

- 3b Je pakt opnieuw 5 balletjes, waarvan er 2 wit zijn. Was zijn nu de a-posteriori waarschijnlijkheden van de hypothesen? Welke verhouding kun jij dus het beste als antwoord kiezen?

Laat Y het aantal witte balletjes na de eerste trekking zijn. Nu zijn de a-priori kansen niet gelijk. Derhalve worden de nieuwe waarschijnlijkheden:

$$P_2(H_i) = P_1(H_i | Y = 2) = \frac{P_1(Y = 2 | H_i) P_1(H_i)}{\sum_{j=1}^3 P_1(Y = 2 | H_j) P_1(H_i)} = \frac{\binom{5}{2} p^3 (1-p)^2 P_1(H_i)}{n},$$

waarbij $m = \binom{5}{2} \left(\frac{1}{4} \frac{3}{4} \frac{3}{4} \times 0.2809533 + \frac{1}{3} \frac{2}{3} \frac{2}{3} \times 0.3507955 + \frac{2}{5} \frac{3}{5} \frac{3}{5} \times 0.3682511 \right)$.

- 3c Als de eerste keer 2 van de 5 balletjes wit en de tweede keer 2 van de 5 balletjes rood zouden zijn geweest, zou je dan nog steeds hetzelfde antwoord geven?

De waarschijnlijkheden na twee updates zijn voor beiden gelijk, zie vraag 4. Maar de antwoorden na de eerste ronde zijn bij beide verschillend. Omdat bij een eerste update met $Y = 2$ de waarschijnlijkheden zijn:

$$\frac{P_0(Y = 2 | H_i)}{\sum_{j=1}^3 P_0(Y = 2 | H_j)} = \frac{\binom{5}{2} p^3 (1-p)^2}{\binom{5}{2} \left(\frac{1}{4} \frac{3}{4} \frac{3}{4} + \frac{1}{3} \frac{2}{3} \frac{2}{3} + \frac{2}{5} \frac{3}{5} \frac{3}{5} \right)}.$$

- 4 Gegeven een proces van Bayesiaans leren met twee hypothesen H en $\bar{H}$ en twee gebeurtenissen E_1 en E_2 . Doet de volgorde van de updates er toe? Dat wil zeggen, is het zo dat de a-posteriori waarschijnlijkheden na een update met eerst E_1 en dan E_2 gelijk zijn aan de a-posteriori waarschijnlijkheden na een update met eerst E_2 en dan E_1 ? Zo ja, bewijs je antwoord, zo nee, geef een tegenvoorbeeld.

We zullen aantonen dat de volgorde er niet toe doet. Laat P_0 de aanvankelijke verdeling van de hypothesen zijn, P_1 de verdeling na een update met E_1 en P_{12} na een update van P_1 met E_2 . P_2 is de verdeling na een update met E_2 en P_{21} na een update van P_2 met E_1 . We tonen aan dat $P_{12} = P_{21}$.

Uit de definitie van voorwaardelijke kans volgt dat voor elke gebeurtenis A geldt dat

$$P_1(A) = P_0(A | E_1) = \frac{P_0(A \cap E_1)}{P_0(E_1)}.$$

Daaruit volgt dat

$$P_{12}(A) = P_1(A | E_2) = \frac{P_1(A \cap E_2)}{P_1(E_2)} = \frac{P_0(A \cap E_2 \cap E_1)}{P_0(E_1) \frac{P_0(E_2 \cap E_1)}{P_0(E_1)}} = \frac{P_0(A \cap E_2 \cap E_1)}{P_0(E_2 \cap E_1)}.$$

Wegens symmetrie geldt volgens dezelfde redenering dat

$$P_{21}(A) = \frac{P_0(A \cap E_2 \cap E_1)}{P_0(E_2 \cap E_1)}.$$

Dat bewijst dat $P_{12} = P_{21}$.

- 5 Hiervan staan er genoeg voorbeelden op de slides.