

Exercitiul 1

Avem un sertar cu 200 de monede identice ca aspect, dintre care i de tipul A , cu probabilitatea aversului la o aruncare de 0.5, iar $200 - i$ de tipul B , cu probabilitatea aversului la o aruncare de 0.8. Alegem aleator o moneda din sertar si o aruncam o data. Care este probabilitatea predictiva a priori sa obtinem avers? Presupunem ca la prima aruncare a acelei monede am obtinut avers. Care este probabilitatea predictiva a posteriori ca la a 2-a aruncare a aceleiasi monede sa obtinem tot avers?

Solutie:

In general, avem o niste date despre care stim ca apartin mai multor distributii din aceeasi familie (aceeasi pdf sau cdf care depinde de un parametru θ). Datele sunt astfel impartite in categorii, in functie de valoarea lui θ care defineste distributia. Ipoteza reprezinta valorile pe care le ia θ , si modul cum sunt distribuite. Verosimilitatea, $p(x|\theta)$ reprezinta pentru fiecare θ , pmf sau cdf a datelor din distributia care depinde de θ . Probabilitatea datelor reprezinta pentru un x din familia de date, $p(x)$ (probabilitatea valorii x din multimea respectiva). Probabilitatea a priori, $p(\theta|x)$, este probabilitatea ca datele cu valoarea x din setul de date sa fie din distributia de parametru θ (sa fie din "categoria/clasa" determinata de θ).

In cazul de fata, ipoteza este "o moneda este in clasa A sau in clasa B". Stim ca $\mathbb{P}(A) = i/200, \mathbb{P}(B) = 1 - i/200$. Pentru clasa A , verosimilitatea este $\mathbb{P}(1|A) = \theta_A = 0.5, \mathbb{P}(0|A) = 1 - \theta_A = 0.5$, iar pentru clasa B , $\mathbb{P}(1|B) = \theta_B = 0.8, \mathbb{P}(0|B) = 1 - \theta_B = 0.2$.

In limbajul "ipoteza/date", avem \mathcal{H} , ipoteza data de parametrul θ care in cazul de fata ia valorile 0.5, 0.8, distribuite Bernoulli, $p(0.5) = i/200, p(0.8) = 1 - i/200$.

Verosimilitatea, si anume $p(\mathcal{D}|\mathcal{H})$ este $p(x|\theta)$. Pentru ca stim ca pentru fiecare θ , x (datele) iau valorile 0 si 1 distribuite Bernoulli de parametru θ , $p(x = 1|\theta) = \theta$.

1. Prima cerinta a problemei este sa obtinem probabilitatea predictiva a priori ca $x = 1$, mai exact $p(x = 1)$. Stim $p(x = 1|\theta_A), p(x = 1|\theta_B), p(x = 0|\theta_A), p(x = 0|\theta_B), p(\theta_A), p(\theta_B)$, deci calculam $p(x = 1)$ cu Legea Probabilitatii Totale. Astfel, $p(x = 1) = p(x = 1|\theta_A) \cdot p(\theta_A) + p(x = 1|\theta_B) \cdot p(\theta_B) = 0.5 \cdot i/200 + (1 - i/200) \cdot 0.8$.
2. A doua cerinta este sa calculam $p(x_2 = 1|x_1 = 1)$, unde x_1 reprezinta selectia unei valori din setul de date, x_2 selectia celei de a doua valori din set. Astfel, $p(x_2 = 1|x_1 = 1) = \frac{p(x_1=1, x_2=1)}{p(x_1=1)}$.

Calculam folosind Legea Probabilitatii Totale: $p(x_1 = 1, x_2 = 1) = p(x_1 = 1, x_2 = 1|\theta_A) \cdot p(\theta_A) + p(x_1 = 1, x_2 = 1|\theta_B) \cdot p(\theta_B)$. Observam urmatorul fapt: daca stim ca x_1, x_2 sunt din aceeasi distributie (corespund aceluasi θ), alegerea lor este independenta una de cealalta. Mai exact, daca ne aflam intr-una dintre distributii si alegem x_1, x_2 nu depinde de x_1 . Astfel, x_1, x_2 sunt independente conditionat la θ_A, θ_B . Deci, $p(x_1 = 1, x_2 = 1) = p(x_1 = 1|\theta_A) \cdot p(x_2 = 1|\theta_A) \cdot p(\theta_A) + p(x_1 = 1|\theta_B) \cdot p(x_2 = 1|\theta_B) \cdot p(\theta_B)$. Cunoastem toate aceste valori din ipoteza si $p(x_1 = 1), p(x_2 = 1)$ a fost calculata anterior.

Exercitiul 2

Presupunem ca avem o moneda cu probabilitate necunoscuta θ a aversului. Aruncam moneda de 110 ori si obtinem i aversuri si $110 - i$ reversuri. Plecand cu a-priori plata pentru θ , aflati pdf a-posteriori.

Solutie:

Fenomenul este de a arunca o moneda cu probabilitatea de avers θ de 110 ori si sa numaram cate aversuri am obtinut. Astfel, datele pe care le avem sunt numere naturale mai mici de 110. Pentru fiecare θ , avem date distribuite Binomial de parametri $n = 110, \theta$.

Stim ca parametrul θ ce defineste distributiile Binomiale(n, θ) este o valoare din $[0, 1]$, distribuita Uniform pe acest interval. Deci, $\theta \sim 1d\theta$.

Deci, stim probabilitatea a-priori, $p(\theta) = 1d\theta$, si verosimilitatile $p(x = k|\theta) = \binom{n}{k} \cdot \theta^k \cdot (1 - \theta)^{n-k}$.

Calculam $p(\theta|x = k) = \frac{p(x=k|\theta) \cdot p(\theta)d\theta}{p(x=k)}$, conform Formulei lui Bayes. Trebuie sa mai calculam $p(x = k)$ cu Legea Probabilitatii Totale pentru cazul continuu deoarece conditionarea se face la valorile lui θ care sunt continue. Astfel, $p(x = k) = \int_{\mathbb{R}} p(x = k|\theta) \cdot p(\theta)d\theta = \int_0^1 p(x = k|\theta)d\theta = \int_0^1 \binom{n}{k} \cdot \theta^k \cdot (1 - \theta)^{n-k}d\theta = \binom{n}{k} \cdot \int_0^1 \theta^k \cdot (1 - \theta)^{n-k}d\theta$. Integrala din dreapta este Beta($k + 1, n + 1 - k$) = $\frac{k! \cdot (n-k)!}{(n+1)!}$ (vezi Curs!), deci $p(x = k) = \frac{n!}{(n-k)! \cdot k!} \cdot \frac{k! \cdot (n-k)!}{(n+1)!} = \frac{1}{n+1}$.

Observam ca datele per total sunt distribuite uniform (avem $n + 1$ valori in setul de date, $0, \dots, n$, fiecare valoare avand probabilitatea de $\frac{1}{n+1}$).

Revenind, $p(\theta|x = k) = (n + 1) \cdot \binom{n}{k} \cdot \theta^k \cdot (1 - \theta)^{n-k}d\theta = (n + 1) \cdot \frac{n!}{(n-k)! \cdot k!} \cdot \theta^k \cdot (1 - \theta)^{n-k}d\theta = \frac{(n+1)!}{(n-k)! \cdot k!} \cdot \theta^k \cdot (1 - \theta)^{n-k}d\theta$, care reprezinta repartitia Beta($k + 1, n + 1 - k$).

Exercitiul 3

Aratati ca repartitia Beta este o repartitie conjugata pentru repartitia geometrica.

Solutie: Stim: probabilitatea a priori este o repartitie Beta(a, b), verosimilitatea este definita de o repartitie geometrica.

Deci, distributia lui $\theta \in [0, 1]$, ipoteza, este data de densitatea $f(\theta) = \frac{(a+b-1)!}{(a-1)! \cdot (b-1)!} \cdot \theta^{a-1} \cdot (1 - \theta)^{b-1}$, cu $a, b \in \mathbb{N}$.

Pentru θ , $p(x = n|\theta) = (1 - \theta)^{n-1} \cdot \theta$.

Calculam $p(\theta|x = n) = \frac{p(x=n|\theta) \cdot p(\theta)d\theta}{p(x=n)}$.

Calculam cu Legea Probabilitatii Totale pentru cazul continuu $p(x = n) = \int_{\mathbb{R}} p(x = k|\theta) \cdot f(\theta)d\theta = \frac{(a+b-1)!}{(a-1)! \cdot (b-1)!} \int_0^1 (1 - \theta)^{n-1} \cdot \theta \cdot \theta^{a-1} \cdot (1 - \theta)^{b-1}d\theta = \frac{(a+b-1)!}{(a-1)! \cdot (b-1)!} \cdot \int_0^1 \theta^a \cdot (1 - \theta)^{n+b-2}d\theta = \frac{(a+b-1)!}{(a-1)! \cdot (b-1)!} \cdot \text{Beta}(a + 1, n + b - 1) = \frac{(a+b-1)!}{(a-1)! \cdot (b-1)!} \cdot \frac{a! \cdot (n+b-2)!}{(a+b+n-1)!}$.

$$\text{Deci, } p(\theta|x = n) = p(x = n|\theta) \cdot p(\theta) d\theta \cdot \frac{(a-1)!(b-1)!}{(a+b-1)!} \cdot \frac{(a+b+n-1)!}{a!(n+b-2)!} = (1-\theta)^{n-1} \cdot \theta \frac{(a+b-1)!}{(a-1)!(b-1)!} \cdot \theta^{a-1} \cdot (1-\theta)^{b-1} \cdot \frac{(a-1)!(b-1)!}{(a+b-1)!} \cdot \frac{(a+b+n-1)!}{a!(n+b-2)!} d\theta = \theta^a \cdot (1-\theta)^{n+b-2} \cdot \frac{(a+b+n-1)!}{a!(n+b-2)!} d\theta.$$

In concluzie, $p(\theta|x = n)$ este repartizata Beta($a + 1, n + b - 1$).