C. R A J S K I (Warszawa)

POSTULAT BAYESA A ENTROPIA

Postulat Bayesa głosi: Jeśli brak jest danych empirycznych o gęstości rozkładu *a priori* nieznanego parametru, a wiadome jest tylko, że zawiera się on w pewnym przedziale skończonym, to należy założyć, że gęstość tego rozkładu jest równomierna. Postulat ten można wyprowadzić w sposób następujący.

Jeśli gęstość rozkładu a priori nie jest znana, to niepewność otrzymania określonego wyniku losowania jest największa. Jeśli za miarę niepewności przyjmiemy entropię, to z braku danych empirycznych o rozkładzie gęstości a priori należy przyjąć taki rozkład, którego entropia jest maksymalna. Tezę tę będziemy nazywali ***postulatem maksymalnej entropii***. Z niego wynika natychmiast postulat Bayesa.

Rzeczywiście, niechaj ξ oznacza nieznany parametr, a H — operator entropii. Wówczas postulat maksymalnej entropii można napisać w postaci

$$\delta H(\xi) = 0. \tag{1}$$

Ponieważ entropię zmiennej losowej ξ określonej w przedziale A o gęstości $f(x)$ określa wzór

$$H(\xi) = -\int_A f(x) \log f(x)\, dx, \tag{2}$$

to jeśli zmienna losowa jest zawarta w przedziale (x_1, x_2), z (1) i (2) wynika, że

$$\delta\Big[-\int_{x_1}^{x_2} f(x)\log f(x)\,dx\Big] = 0.$$

Przyjęto tu dla uproszczenia logarytmy naturalne. Poszukiwana funkcja $f(x)$, jako funkcja gęstości, powinna spełniać warunek dodatkowy

$$\int_{x_1}^{x_2} f(x)\,dx = 1.$$

Rozwiązaniem tych dwóch równań jest

$$f(x) = 1/(x_2 - x_1), \tag{3}$$

co zresztą jest konsekwencją założenia, że określona przez (2) funkcja entropii powinna osiągać wartość maksymalną dla rozkładu równomiernego. Dowiedliśmy zatem, że reguła Bayesa wynika z postulatu maksymalnej entropii.

Jeśli długość przedziału zmienności ξ rośnie bez granic, to gęstość dana wzorem (3) dąży do granicy, która nie jest gęstością, wobec czego sama reguła Bayesa staje się bezużyteczna. Jeśli jednakże wyjść z postulatu maksymalnej entropii, to można usunąć tę trudność stawiając pewne, bardzo słabe zresztą, warunki dodatkowe.

Dla przedziału $(0, \infty)$ wystarczy przyjąć, że wartość oczekiwana a zmiennej ξ jest skończona. W tym przypadku otrzymamy zatem następujący układ równań:

$$\delta\left[-\int_0^\infty f(x)\log f(x)\,dx\right] = 0, \tag{4}$$

$$\int_0^\infty f(x)\,dx = 1, \tag{5}$$

$$\int_0^\infty x f(x)\,dx = a < \infty. \tag{6}$$

Stąd metodą współczynników Lagrange'a znajdujemy, że

$$\delta\int_0^\infty [-f(x)\log f(x) + \lambda x f(x) + \mu f(x)]\,dx = 0,$$

i po zróżniczkowaniu względem f dochodzimy do równania

$$-\log f(x) - 1 + \lambda x + \mu = 0,$$

a stąd

$$f(x) = e^{\lambda x + \mu - 1}. \tag{7}$$

Podstawiając (7) do (5) i (6) otrzymujemy $e^{\mu-1} = -\lambda$ oraz $e^{\mu-1} = a\lambda^2$. Po rozwiązaniu i wstawieniu do (7) znajdujemy ostatecznie

$$f(x) = \frac{1}{a}e^{-x/a}. \tag{8}$$

W przypadku przedziału $(-\infty, \infty)$ warunek (6) jest niedostateczny, bo wówczas entropia nie zależy od wartości oczekiwanej, tzn. że dla każdej liczby rzeczywistej a zachodzi równość $H(\xi + a) = H(\xi)$. Można jednakże znaleźć i w tym przypadku rozkład o maksymalnej entropii, stawiając jako warunek dodatkowy istnienie skończonego odchylenia średniego. Ponieważ liczba rozwiązań różniących się tylko wartością oczekiwaną jest nieskończona, wystarczy znaleźć jedno z nich, np. takie, dla którego wartość oczekiwana jest równa zeru. Wówczas układ równań określający poszukiwaną gęstość rozkładu ma postać następującą:

$$\delta\left[-\int_{-\infty}^{\infty} f(x)\log f(x)\,dx\right] = 0, \tag{9}$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1, \tag{10}$$

$$\int_{-\infty}^{\infty} x^2 f(x)\,dx = \sigma^2 < \infty. \tag{11}$$

Stosując ponownie metodę współczynników Lagrange'a otrzymujemy

$$\delta \int_{-\infty}^{\infty} [-f(x)\log f(x) + \lambda x^2 f(x) + \mu f(x)]\,dx = 0$$

i po zróżniczkowaniu względem f

$$-\log f(x) - 1 + \lambda x^2 + \mu = 0,$$

a stąd

$$f(x) = e^{\lambda x^2 + \mu - 1}. \tag{12}$$

Wstawiając (12) do (10) i (11) otrzymujemy

$$e^{\mu-1} = \sqrt{-\frac{\lambda}{\pi}}, \quad \frac{-e^{\mu-1}}{2\lambda}\sqrt{-\frac{\pi}{\lambda}} = \sigma^2,$$

a stąd

$$\lambda = -1/2\sigma^2, \quad e^{\mu-1} = 1/\sigma\sqrt{2\pi}$$

i ostatecznie

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}.$$

W ten sposób dowiedliśmy, że postulat maksymalnej entropii ma szersze znaczenie niż postulat Bayesa, ponieważ pozwala on na znalezienie funkcji gęstości *a priori* również w przypadkach przedziału półnieskończonego i nieskończonego.

Wszystkie przytoczone wzory zostały zaczerpnięte z podstawowej pracy Shannona i wyłącznym zadaniem niniejszego artykułu jest wykazanie ich użyteczności w problematyce postulatu Bayesa.

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK

Praca wpłynęła 19. 12. 1956

Ч. РАЙСКИЙ (Варшава)

ПОСТУЛАТ БЕЙЕСА И ЭНТРОПИЯ

РЕЗЮМЕ

Доказывается, что постулат Бейеса о равномерном распределении плотности априори можно вывести из общего постулата максимума энтропии.

Кроме того в работе вычислены распределения случайного переменного, удовлетворяющего этому постулату, в случае, когда переменное заключено в интервале $(0, \infty)$ или $(-\infty, \infty)$. В первом случае предполагается дополнительно существование первого момента; искомое распределение есть распределение типа гамма. Во втором случае предполагается существование второго момента и искомое распределение нормально.

C. RAJSKI (Warszawa)

THE BAYES POSTULATE AND ENTROPY

SUMMARY

It is proved in the paper that the Bayes postulate on the uniform distribution of a priori density can be deduced from the more general postulate of maximum entropy.

Moreover the author determines the distributions of the random variable satisfying this postulate in the case where the random variable is contained in the intervals $(0, \infty)$ and $(-\infty, \infty)$. In the first case we assume the additional assumption of the existence of the first moment. In the second case the existence of the second moment is assumed and then the desired distribution is normal.
