

PETRESCU Mary-Violeta

**DISTRIBUȚIA DE PROBABILITATE A UNEI
VARIABLE ALEATOARE**

Craiova
Mai, 2025

Copyright © 2025, Autor
Tehnoredactare computerizată: PETRESCU Mary-Violeta
Corectura aparține autorului

Responsabilitatea pentru conținutul lucrării revine în exclusivitate autorului
Studiu științific

Publicat: Mai 2025
<http://www.edituraelse.ro/ped/ped.htm>

DISTRIBUȚIA DE PROBABILITATE A UNEI VARIABILE ALEATOARE

Rezolvarea problemelor privind probabilitatea anumitor evenimente desfășurate în anumite condiții a început încă din primele decenii ale secolului al XVIII-lea.

Una din teoremele fundamentale ale teoriei probabilităților a fost enunțată de matematicianul Thomas Bayes.

Teorema fost citită de către Richard Price la Royal Society (1763) în „Essay Towards Solving a Problem in the Doctrine of Chances” după moartea lui Bayes în 1761, fiind publicată în Philosophical Transactions of the Royal Society of London (1764).

Odată cu dezvoltarea metodelor de analiză statistice, teorema lui Bayes este întâlnită în diverse studii, termenul „bayesian” fiind folosit încă din 1950.

Determinarea probabilităților unui eveniment E care aparține unui sistem complet de evenimente notat cu Ω definește noțiunea de probabilitate totală.

Probabilitatea de apariție a unuia din evenimentele $E_1, E_2, \dots, E_n$ care aparțin sistemului complet de evenimente Ω este dată de următoarea relație, care este formula probabilității totale:

$$P(E) = P(E_1) \cdot P(E/E_1) + P(E_2) \cdot P(E/E_2) + \dots + P(E_n) \cdot P(E/E_n) \quad (1)$$

Bayes a ajuns la concluzia că dacă inversează formula probabilității totale, atunci probabilitatea ca apariția evenimentului E să se datoreze apariției unui alt eveniment E_k ținând cont de faptul, că sunt cunoscute condițiile în care probabilitatea de apariție a evenimentului E și probabilitățile de apariție ale unor evenimente E_i , care condiționează apariția evenimentului E .

În ceea ce urmează vom defini formula lui Bayes.

Se dă un sistem complet de evenimente Ω și presupunând că probabilitatea unui eveniment E_i îndeplinește condiția $P(E_i) > 0$, atunci formula lui Bayes va fi definită astfel:

$$P(E_k/E) = \frac{P(E_k) \cdot P(E/E_k)}{P(E)} = \frac{P(E_k) \cdot P(E/E_k)}{\sum_{i=1}^n P(E_i) \cdot P(E/E_i)} \quad (2)$$

În relația (2) termenul $P(E_k/E)$ reprezintă probabilitatea aposteriorică, iar probabilitățile $P(E_k)$ reprezintă probabilitățile apriorice.

În continuare, va fi prezentată o alta formă a formulei lui Bayes în care probabilitatea aposteriorică este calculată ca produs între probabilitatea apriorică și ponderea pe care o are probabilitatea evenimentului E condiționată de evenimentul E_k în probabilitatea totală a evenimentului E :

$$P(E_k/E) = y_k \cdot P(E_k) \quad (3)$$

unde $y_k = \frac{P(E/E_k)}{P(E)}$ reprezintă ponderea pe care o are probabilitatea evenimentului E

condiționată de evenimentul E_k în probabilitatea totală a evenimentului E .

Până în acest moment teorema lui Bayes a fost prezentată pentru cazul unidimensional, caz în care s-au folosit termenii eveniment, probabilități și probabilități condiționate.

În analiza cantitativă economică sunt utilizate două concepte fundamentale cum ar fi variabila aleatoare, și respectiv distribuție de probabilitate a unei variabile aleatoare. Variabila aleatoare poate fi o variabilă reală cu valori care sunt dependente de rezultatele unui experiment aleator.

O variabilă aleatoare notată cu X poate fi definită ca o funcție, în condițiile în care sunt cunoscute următoarele elemente:

- ω - un rezultat posibil al unui experiment aleator;
- Ω - spațiul de selecție format din mulțimea rezultatelor experimentului aleator;
- $(\Omega, \mathcal{F})$ - câmpul evenimentelor peste care variabila aleatoare este definită;
- $(\Omega, \mathcal{F}, \mathbf{P})$ - câmpul de probabilitate.

Fiind date elementele, funcția va avea următoarea formă:

$$X : \Omega \rightarrow \mathbf{R}, (X(\omega) = x \in \mathbf{R}) \quad (4)$$

Forma funcției rezultată este aplicată submulțimii de evenimente

$$\{\omega | X(\omega) \leq x, (\forall) x \in \mathbf{R}\} \quad (5)$$

care va fi element al câmpului de evenimente, respectiv

$$\{\omega | X(\omega) \leq x, (\forall) x \in \mathcal{F}\} \quad (6)$$

Distribuția de probabilitate a variabilei aleatoare reprezintă asocierea cu probabilități a mulțimii valorilor unei variabile aleatoare X , iar modalitatea de specificare a distribuției se realizează cu ajutorul unei funcții reale care poartă numele de funcție de repartiție, notată cu $F_X(\cdot)$.

O altă modalitate de specificare a distribuției de probabilitate a variabilei aleatoare este realizată în cazul în care funcția de repartiție este diferențiabilă, adică:

$$f_X(x) = \frac{dF_X(x)}{dx} \quad (7)$$

Funcția din relația (4) este numită densitate de probabilitate.

Prin urmare funcția de repartiție va fi de forma:

$$F_X(x) = \int_{-\infty}^x f_X(x) dx \quad (8)$$

În cele ce urmează va fi prezentată teorema lui Bayes pentru cazul multidimensional care păstrează aceleași principii ca în cazul unidimensional.

Pentru cazul multidimensional vor fi introduși termenii de vectori aleatori, densități de probabilitate și densități marginale de probabilitate, care vor fi prezentați în cele ce urmează.

Datorită faptului că în analiza multidimensională participă mai multe variabile aleatoare notate cu $X_1, X_2, \dots, X_n$ vom presupune că acestea pot fi elementele unui vector numit vector aleator. Orice obiect definit prin două sau mai multe caracteristici poate să aibă asociat o variabilă aleatoare vectorială.

Vectorul aleator $X = (X_1, X_2, \dots, X_n)^t$ se poate defini astfel:

$$X: \mathbf{S}^n \rightarrow \mathbf{R}^n, \mathbf{S}^n \text{ fiind spațiul de selecție n-dimensional} \quad (9)$$

Funcția de repartiție a vectorului aleator X este definită astfel:

$$F_X: \mathbf{R}^n \rightarrow [0,1] \quad (10)$$

$$F_X(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

Densitatea de probabilitate n-dimensională a vectorului aleator X este o funcție ca în relația de mai jos:

$$f_X(x_1, x_2, \dots, x_n) = \frac{\partial^n F_X(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n} \quad (11)$$

care este definită numai atunci când funcția de repartiție $F_X(\cdot, \dots, \cdot)$ este diferențiabilă în raport cu toate argumentele.

Prin urmare, densitatea de probabilitate n-dimensională a vectorului aleator va avea următoarea formă:

$$F_X(x_1, x_2, \dots, x_n) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} f_X(u_1, u_2, \dots, u_n) du_n du_{n-1} \dots du_2 du_1 \quad (12)$$

Cunoscând densitățile multidimensionale de probabilitate a vectorului aleator, se poate defini densitatea de probabilitate a oricărei variabile aleatoare X_i care va purta numele de densitate marginală a vectorului aleator X .

Densitatea marginală a vectorului aleator va avea următoarea formă:

$$g_{X_i}(x_i) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_X(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) dx_n \dots dx_{i+1} dx_{i-1} \dots dx_1 \quad (13)$$

La rândulei, funcția de repartiție a vectorului aleator se poate defini și va purta numele de funcție de repartiție marginală a vectorului aleator, astfel:

$$G_{X_i}(x_i) = F_X(\infty, \dots, \infty, x_i, \dots, \infty) \quad (14)$$

Pentru exprimarea teoremei lui Bayes se consideră:

- vectorii aleatori X și Y cu dimensiunile m , respectiv n , iar densitatea de probabilitate multidimensională notată cu $f(.)$ definește distribuția comună a vectorilor;
- probabilitatea apriorică a densității de probabilitate a lui X sau densitatea marginală pentru vectorul aleator X , notată cu $g_1(x)$;
- probabilitatea apriorică a densității de probabilitate a lui Y sau densitatea marginală pentru vectorul aleator Y , notată cu $g_2(y)$;
- probabilitatea apriorică $h_1(x/y)$ care este densitatea de probabilitate a lui X , condiționată fiind de vectorul aleator Y care ia valorile reprezentate de vectorul y ;
- probabilitatea aposteriorică $h_2(y/x)$ care este densitatea de probabilitate a lui Y , condiționată fiind de vectorul aleator X care ia valorile reprezentate de vectorul x .

Teorema lui Bayes pentru cazul multidimensional este exprimată pentru $(\forall) y \in \mathfrak{R}^n$ cu condiția ca $g_2(y) > 0$ și $(\forall) x \in \mathfrak{R}^m$, astfel:

$$h_1(x/y) = \frac{h_2(y/x) \cdot g_1(x)}{g_2(y)} = \frac{h_2(y/x) \cdot g_1(x)}{\int_{\mathfrak{R}^m} h_2(y/\xi) g_1(\xi) d\xi} \quad (15)$$

BIBLIOGRAFIE:

1. Sarkar, S. & Sriram, R.S., 2001. Bayesian Models, 47(11), pp.1457-1475.
2. https://en.wikipedia.org/wiki/Bayesian_probability.
3. Ben-Gal I., Bayesian Networks, in Ruggeri F., Faltin F. & Kenett R., Encyclopedia of Statistics in Quality & Reliability, Wiley & Sons (2007).