Rachunek Prawdopodobieństwa MAP1181

Wydział Matematyki, Matematyka Stosowana

Projekt - losowy przedział

Opracowanie: Weronika Nitka

Zadanie:

Z odcinka $[0,1]$ wylosowano niezależnie punkty A, B i C wg rozkładu jednostajnego. Jakie jest prawdopodobieństwo zdarzenia, że punkt C należy do przedziału o końcach A i B? Odpowiedź uzasadnij.

Rozwiązanie:

A, B i C to niezależne zmienne losowe, których realizacje to trzy punkty wylosowane z odcinka $[0,1]$. Dla ustalonych punktów A i B prawdopodobieństwo, że punkt C należy do przedziału o końcach A, B możemy wyrazić za pomocą modelu prawdopodobieństwa geometrycznego. Jest to możliwe dlatego, że punkt C losujemy według rozkładu jednostajnego, tj. w każdym punkcie gęstość rozkładu zmiennej losowej jest identyczna. Szukana przez nas wartość $P\left(C \in [A,B]\right)$ będzie wyrażała się stosunkiem długości przedziału $[A,B]$ (lub przedziału $[B,A]$ w przypadku, gdy $B < A$) do długości przedziału $[0,1]$ równej 1. Stosunek ten będzie równy długości odcinka $|AB|$.

Ponieważ punkty A i B nie są ustalone, ale również są losowane z odcinka $[0,1]$ według rozkładu jednostajnego, $A \sim \mathcal{U}(0,1)$, $B \sim \mathcal{U}(0,1)$, odległość $|AB|$ nie jest znana. Szukane prawdopodobieństwo wyliczyć można techniką warunkowej wartości oczekiwanej:

$$
\begin{aligned}
&P(C \text{ leży pomiędzy } A, B) = \mathrm{E}P(C \text{ leży pomiędzy } A, B|A, B) = \\
&= \mathrm{E}\left[P(C \text{ leży pomiędzy } A, B|A, B)\Big|_{\substack{a=A\\ b=B}} \right] \overset{(*)}{=} \mathrm{E}\left[P(C \text{ leży pomiędzy } a, b)\Big|_{\substack{a=A\\ b=B}} \right] = \\
&= \mathrm{E}\left[\frac{|b-a|}{1}\Big|_{\substack{a=A\\ b=B}} \right] = \mathrm{E}|B-A|,
\end{aligned}
$$

gdzie w przejściu oznaczonym $(*)$ korzystamy z niezależności zmiennych losowych A, B i C.

Punkty A, B i C losujemy według rozkładu jednostajnego $\mathcal{U}(0,1)$, którego dystrybuanta wyraża się wzorem

$$
F(x) = \begin{cases} 0, \ x \leqslant 0 \\ x, \ 0 < x \leqslant 1 \\ 1, \ x > 1, \end{cases} \tag{1}
$$

zaś gęstość

$$
f(x) = \begin{cases} 0, \ x \notin [0,1] \\ 1, \ x \in [0,1]. \end{cases} \tag{2}
$$

Rozważymy teraz wektor losowy (A,B). Zmienne losowe A i B są niezależne o gęstościach $f_A \equiv f_B \equiv f$. Przez x oznaczamy wartość punktu A, zaś przez y - wartość punktu B. Gęstość rozkładu łącznego zmiennych losowych A i B będzie miała postać

$$
f_{A,B}(x,y) = f_A(x)f_B(y) = f(x)f(y) = \begin{cases} 1, \ x \in [0,1], \ y \in [0,1] \\ 0 \text{ w pozostałych punktach.} \end{cases} \tag{3}
$$

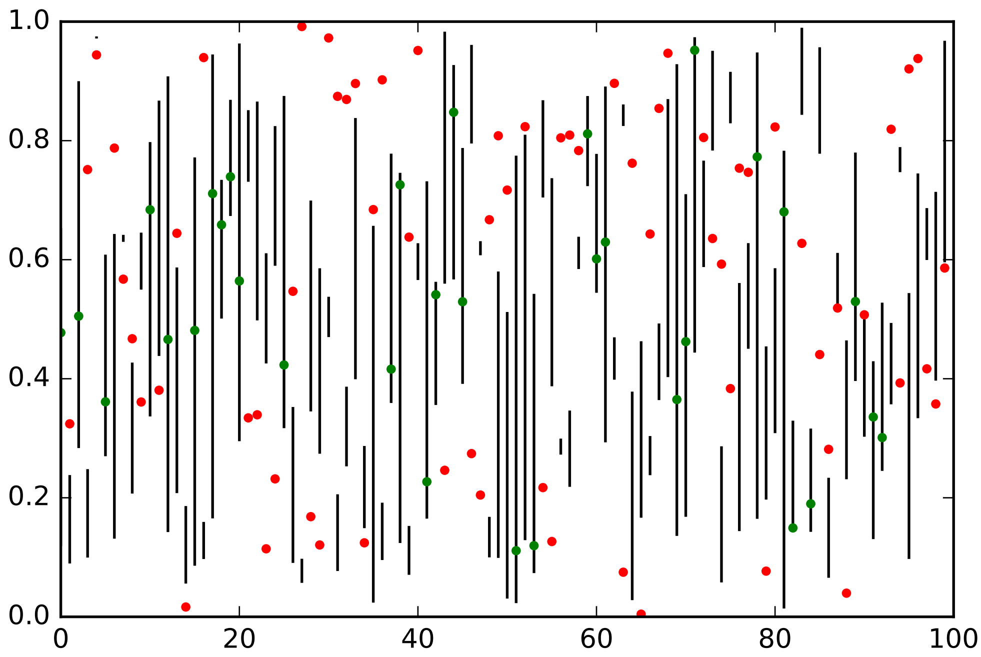
Wartość oczekiwana zmiennej losowej $|B-A|$ wyraża się wzorem

$$\mathrm{E}|B-A| = \mathrm{E}g(A,B) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{A,B}(x,y)dxdy, \tag{4}$$

gdzie $g(x,y) = |y-x|$. Otrzymujemy

$$\begin{aligned}
\mathrm{E}|B-A| &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{A,B}(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |y-x| f_{A,B}(x,y)dxdy = \int_0^1\int_0^1 |y-x|dxdy = \\
&= \int_0^1\int_0^y (y-x)dxdy + \int_0^1\int_y^1 (x-y)dxdy = \int_0^1\int_0^y y\ dxdy - \int_0^1\int_0^y x\ dxdy + \int_0^1\int_y^1 x\ dxdy + \\
&- \int_0^1\int_y^1 y\ dxdy = \int_0^1 y^2\ dy - \int_0^1 \frac{y^2}{2}\ dy + \int_0^1 \frac{1}{2}\ dy - \int_0^1 \frac{y^2}{2}\ dy - \int_0^1 y\ dy + \int_0^1 y^2\ dy = \\
&= \int_0^1 y^2\ dy + \int_0^1 \frac{1}{2}\ dy - \int_0^1 y\ dy = \frac{1}{3} + \frac{1}{2} - \frac{1}{2} = \frac{1}{3}
\end{aligned}$$

Wartość oczekiwana odległości pomiędzy punktami A i B wynosi $\frac{1}{3}$, zatem, zgodnie z wcześniejszym spostrzeżeniem, prawdopodobieństwo, że punkt C będzie należał do przedziału o końcach A i B również wynosi $\frac{1}{3}$. Przeprowadzenie symulacji komputerowej losowania takich trzech punktów potwierdza otrzymany wynik - z przedstawionej na poniższym wykresie próby stu wylosowanych trójek A, B, C, 32 z nich zawiera punkt C (zielone punkty na wykresie) w przedziale o końcach A i B (czarne linie), zaś pozostałych 68 daje punkt C poza tym przedziałem (czerwone punkty na wykresie). Symulacje na większych próbach dają jeszcze lepsze przybliżenie teoretycznej częstości $\frac{1}{3}$.



Rysunek 1: Symulacja losowania stu trójek punktów A, B, C. Otrzymana częstość wystepowania punktu C pomiędzy punktami A i B wynosi 0.32, co stanowi dobre przybliżenie wyliczonej częstości teoretycznej $\frac{1}{3}$.