RACHUNEK PRAWDOPODOBIEŃSTWA

ZMIENNE LOSOWE

DWUWYMIAROWE

Zmienne losowe dwuwymiarowe

Często w doświadczeniu losowym mamy do czynienia z kilkoma wielkościami losowymi, które w jakiś sposób są ze sobą powiązane i do opisu eksperymentu niewystarczające jest podanie jedynie rozkładów pojedynczych zmiennych.

Najpierw rozważymy przypadek dwóch zmiennych, aby poznać różnice między przypadkiem jednowymiarowym a dwuwymiarowym. Następnie uzyskane wyniki rozszerzymy na przypadek większej liczby zmiennych.

Definicja.

Mówimy, że (X, Y) jest dwuwymiarową zmienną losową, jeżeli X i Y są dwiema zmiennymi losowymi zdefiniowanymi na tej samej przestrzeni zdarzeń elementarnych Ω

$$(X, Y) : \Omega \longrightarrow R^2.$$

oraz

$$\{\omega \in \Omega : X(\omega) \leqslant x, Y(\omega) \leqslant y\} \in \mathcal{F} \quad \forall (x, y) \in \mathcal{R}^2$$

jest zdarzeniem losowym.

Analogicznie jak w przypadku jednowymiarowym, zdefiniujemy pojęcie dystrybuanty dwuwymiarowej zmiennych losowych X i Y.

Definicja.

Dystrybuantą dwuwymiarową zmiennych losowych X i Y nazywamy taką funkcję dwóch zmiennych F, że dla dowolnego $(x, y) \in \mathcal{R}^2$

$$F(x, y) = P(X \leqslant x, Y \leqslant y) = P(\{\omega \in \mathcal{S} : X(\omega) \leqslant x, Y(\omega) \leqslant y\}).$$

Innymi słowy, $F(x, y)$ jest prawdopodobieństwem zdarzenia takim, że dwuwymiarowa zmienna losowa (X, Y) przyjmuje wartości ze zbioru $\{(s, t) \in R^2 : s \leqslant x, t \leqslant y\}$.

Dystrybuanta dwuwymiarowa spełnia podobne warunki konieczne i dostateczne jak dystrybuanta jednowymiarowa.

Własności dystrybuanty dwuwymiarowej

1. $\lim_{x\to-\infty} F(x,y) = 0$, $\lim_{y\to-\infty} F(x,y) = 0$, oraz $\lim_{x,y\to\infty} F(x,y) = 1$.
2. Dystrybuanta dwuwymiarowa jest funkcją niemalejącą ze względu na każdą ze zmiennych, tzn. $\forall x\ F(x,y)$ oraz $\forall y\ F(x,y)$ są, odpowiednio, funkcjami niemalejącymi ze względu na y i x.
3. Dystrybuanta dwuwymiarowa jest funkcją prawostronnie ciągłą ze względu na każdą ze zmiennych.
4. Dla dowolnych $x_1 < x_2$ oraz $y_1 < y_2$

$$F(x_1,y_1) + F(x_2,y_2) - F(x_1,y_2) - F(x_2,y_1) \geqslant 0.$$

Twierdzenie 1.
Warunki 1-4 są warunkami koniecznymi i wystarczającymi na to, aby dwuwymiarowa funkcja F była dystrybuantą pewnej zmiennej losowej.

Przykład. *Pokazać, że funkcja*

$$F(x, y) = \begin{cases} 1 & \text{dla} \quad x + y \geqslant 1, \ x \geqslant 0, \ y \geqslant 0 \\ 0 & \quad \text{w przeciwnym wypadku} \end{cases}$$

spełnia warunki 1-3, ale nie spełnia warunku 4, zatem F nie jest dystrybuantą żadnej zmiennej losowej.

Rozwiązanie.

Istotnie, dla $x_1 = y_1 = 0$ oraz $x_2 = y_2 = 1$ mamy
$F(0, 0) + F(1, 1) - F(1, 0) - F(0, 1) = -1.$

Dwuwymiarowe zmienne losowe typu dyskretnego

W przypadku, gdy obie zmienne losowe X i Y są typu dyskretnego z nośnikami $\mathcal{X} = \{x_1, x_2, \dots\}$ oraz $\mathcal{Y} = \{y_1, y_2, \dots\}$, można zdefiniować funkcję prawdopodobieństwa rozkładu łącznego.

Definicja.

Funkcję $p(x, y)$ zdefiniowaną wzorem

$$p(x, y) = P(X = x, Y = y) \quad \forall (x, y) \in R^2.$$

nazywamy **funkcją prawdopodobieństwa rozkładu łącznego** dwuwymiarowej zmiennej losowej (X, Y).

Własności funkcji prawdopodobieństwa $p(x, y)$

1. $p(x, y) \geqslant 0, \ \forall (x, y) \in R^2.$
2. $\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x, y) = 1.$

Twierdzenie

Warunki 1-2 są warunkami koniecznymi i dostatecznymi na to, aby funkcja $p(x, y)$ była funkcją prawdopodobieństwa dwuwymiarowej zmiennej losowej typu dyskretnego.

Dla wszystkich $x \in \mathcal{X}$ mamy

$$\{X = x\} = \bigcup_{y \in \mathcal{Y}} \{X = x, Y = y\},$$

pozwala nam to zdefiniować funkcję prawdopodobieństwa zmiennej losowej X zwaną funkcją prawdopodobieństwa **rozkładu brzegowego**.

Definicja.

Funkcje prawdopodobieństwa ***rozkładów brzegowych*** zmiennych losowych X i Y, oznaczonych jako $p_X(x)$ and $p_Y(y)$, są równe odpowiednio

$$p_X(x) = P(X = x) = \sum_{y \in \mathcal{Y}} p(x, y) \quad p_Y(y) = P(Y = y) = \sum_{x \in \mathcal{X}} p(x, y).$$

Przykład. *Towarzystwo ubezpieczeniowe świadczy ubezpieczenia komunikacyjne oraz ubezpieczenia mieszkań. Dla każdego typu ubezpieczenia są stosowane zniżki. Niech X będzie zniżką na polisę komunikacyjną (w zł) a Y zniżką na polisę mieszkaniową udzielaną klientowi.*
Łączny rozkład prawdopodobieństwa zmiennej losowej (X, Y) dany jest poprzez funkcję prawdopodobieństwa zdefiniowaną w tabeli poniżej:

	Y			
X	0	100	200	300
0	0.10	0.15	0.15	0.05
100	0.05	0.05	0.10	0.10
250	0.05	0.05	0.05	0.10

Znaleźć rozkłady brzegowe zmiennych losowych X i Y.
Obliczyć prawdopodobieństwo zdarzenia $\{Y \geqslant 200\}$.

Rozwiązanie.
Rozkłady brzegowe zmiennych losowych X i Y znajdujemy, sumując prawdopodobieństwa w wierszach i w kolumnach, odpowiednio.

	Y				
X	0	100	200	300	$p_X(x)$
0	0.10	0.15	0.15	0.05	0.45
100	0.05	0.05	0.10	0.10	0.30
250	0.05	0.05	0.05	0.10	0.25
$p_Y(y)$	0.20	0.25	0.30	0.25	

Z tabeli można odczytać, że

$$P(Y \geqslant 200) = p_Y(200) + p_Y(300) = 0.30 + 0.25 = 0.55.$$

Dwuwymiarowe zmienne losowe typu ciągłego

Niech (X, Y) będzie parą zmiennych losowych o rozkładach ciągłych przyjmującą wartości ze zbioru $\mathcal{X} \times \mathcal{Y}$, gdzie $\mathcal{X}$ i $\mathcal{Y}$ są, odpowiednio, nośnikami zmiennych losowych X i Y.

Podobnie jak w przypadku jednowymiarowym, możemy zdefiniować funkcję gęstości prawdopodobieństwa, tym razem rozkładu łącznego, która pozwoli nam wyznaczyć prawdopodobieństwo zdarzenia, że zmienna losowa (X, Y) przyjmie wartości z dwuwymiarowego zbioru A, za pomocą całki z funkcji gęstości po zbiorze A.

Definicja.

Niech X i Y będą dwiema zmiennymi losowymi typu ciągłego. Wówczas nieujemna funkcja $f(x, y)$ jest ***gęstością prawdopodobieństwa rozkładu łącznego*** X i Y, jeżeli dla dowolnego zdarzenia losowego $A \subset \mathcal{R}^2$

$$P\{(X, Y) \in A\} = \iint_A f(x, y)\mathrm{d}x\,\mathrm{d}y.$$

W szczególności, dla dowolnych $(x, y) \in R^2$ otrzymujemy dystrybuantę łączną $F(x, y)$

$$F(x, y) = P\{(X, Y) \in (-\infty, x] \times (-\infty, y]\} = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u, v)\mathrm{d}u\,\mathrm{d}v.$$

Różniczkując dystrybuantę w punktach ciągłości gęstości f mamy

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y).$$

Własności gęstości dwuwymiarowej

1. $f(x,y) \geqslant 0, \ \forall (x,y) \in R^2$.
2. $\iint\limits_{R^2} f(x,y) \mathrm{d}x \, \mathrm{d}y = 1$.

Twierdzenie

Warunki 1-2 są warunkami koniecznymi i wystarczającymi na to, aby funkcja $f(x,y)$ była łączną gęstością prawdopodobieństwa pewnej zmiennej losowej typu ciągłego.

Jeżeli zmienne losowe X i Y mają łączny rozkład typu ciągłego, to rozkłady brzegowe są także typu ciągłego a ich gęstości można uzyskać w następujące sposób:

$$\begin{aligned} P\{X \in A\} &= P\{X \in A, Y \in (-\infty, \infty)\} = \int\limits_A \int\limits_{-\infty}^{\infty} f(x, y) \mathrm{d}x \, \mathrm{d}y = \\ &= \int\limits_A f_X(x) \mathrm{d}x, \end{aligned}$$

gdzie

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}y$$

jest gęstością prawdopodobieństwa zmiennej losowej X.

Definicja.

Gęstościami brzegowymi zmiennych losowych X i Y, oznaczonymi odpowiednio jako $f_X(x)$ oraz $f_Y(y)$, są

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y \quad \text{dla} \quad -\infty < x < \infty,$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}x \quad \text{dla} \quad -\infty < y < \infty.$$

Przykład. *Pewna sieć typu fast-food posiada okienka dla pieszych oraz okienka dla kierowców. W losowo wybranym dniu, niech X oznacza proporcję czasu, kiedy okienko dla zmotoryzowanych jest zajęte (co najmniej jeden klient jest obsługiwany lub czeka na obsługę) a Y oznacza proporcję czasu zajętości okienka dla niezmotoryzowanych. Nośnikiem rozkładu łącznego (X, Y) jest kwadrat $D = \{(x, y) : 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1\}$.*
Załóżmy, że gęstość dwuwymiarowa (X, Y) dana jest przez

$$f(x, y) = \begin{cases} \frac{6}{5}(x + y^2) & \text{jeżeli} \quad 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1 \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

(a) *Sprawdzić, czy $f(x, y)$ jest gęstością prawdopodobieństwa.*
(b) *Obliczyć prawdopodobieństwo, że oba okienka nie są zajęte jednocześnie dłużej niż przez 25% czasu.*
(c) *Wyznaczyć gęstości brzegowe oraz obliczyć $P(0,25 \leqslant Y \leqslant 0,75)$.*

Rozwiązanie.

(a) Zauważmy, że $f(x,y) \geqslant 0$. Ponadto

$$\begin{aligned}
\iint_{R^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_0^1\int_0^1 f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_0^1\int_0^1 \frac{6}{5}(x+y^2)\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_0^1\int_0^1 \frac{6}{5}x\,\mathrm{d}x\,\mathrm{d}y + \int_0^1\int_0^1 \frac{6}{5}y^2\,\mathrm{d}x\,\mathrm{d}y = \\
&= \int_0^1 \frac{6}{5}x\,\mathrm{d}x\,\mathrm{d}y + \int_0^1 \frac{6}{5}y^2\,\mathrm{d}x\,\mathrm{d}y = \\
&= \frac{6}{10} + \frac{6}{15} = 1.
\end{aligned}$$

(b) Prawdopodobieństwo tego, że oba okienka nie są zajęte jednocześnie dłużej niż przez 25% czasu wynosi

$$P(0 \leqslant X \leqslant 0.25, 0 \leqslant Y \leqslant 0.25) =$$

$$= \int_0^{0.25}\int_0^{0.25} \frac{6}{5}(x+y^2)\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \frac{6}{5}\int_0^{0.25}\int_0^{0.25} x\,\mathrm{d}x\,\mathrm{d}y + \frac{6}{5}\int_0^{0.25}\int_0^{0.25} y^2\,\mathrm{d}x\,\mathrm{d}y =$$

$$= \frac{6}{20}\cdot\frac{x^2}{2}\Big|_0^{0.25} + \frac{6}{20}\cdot\frac{y^3}{3}\Big|_0^{0.25} = \frac{7}{640} = 0.0109.$$

(c) Gęstość rozkładu brzegowego czasu zajętości okienka dla zmotoryzowanych X (bez odwoływania się do czasu zajętości okienka dla niezmotoryzowanych) dla $0 \leqslant x \leqslant 1$ jest równa

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\mathrm{d}y = \int_0^1 \frac{6}{5}(x+y^2)\,\mathrm{d}y = \frac{6}{5}x + \frac{2}{5}.$$

Zatem

$$f_X(x) = \begin{cases} \frac{6}{5}x + \frac{2}{5} & \text{dla} \quad 0 \leqslant x \leqslant 1 \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

Podobnie otrzymujemy rozkład brzegowy zmiennej Y.

$$f_Y(y) = \begin{cases} \frac{6}{5}y^2 + \frac{3}{5} & \text{dla} \quad 0 \leqslant y \leqslant 1 \\ 0 & \text{w przeciwnym przypadku.} \end{cases}$$

Zatem

$$P(0.25 \leqslant Y \leqslant 0.75) = \int_{0.25}^{0.75} f_Y(y)\,\mathrm{d}y = \frac{37}{80} = 0.4625.$$

Niezależność zmiennych losowych

W wielu sytuacjach, informacja o wartości przyjmowanej przez jedną ze zmiennych daje informację o możliwych wartościach drugiej ze zmiennych. Mówimy wówczas, że zmienne są zależne.

Już wcześniej dowiedzieliśmy się, że niektóre zdarzenia losowe są wzajemnie niezależne. Używając podobnych argumentów możemy zdefiniować niezależne zmienne losowe.

Definicja.

Zmienne losowe X i Y są niezależne wtedy i tylko wtedy, gdy

$$F(x, y) = F_X(x) \cdot F_Y(y) \quad \text{dla dowolnych} \quad x, y \in \mathcal{R}.$$

W praktyce, warunek niezależności podany w definicji, nie jest zbyt wygodny do sprawdzenia. Szczególnie w przypadku, gdy zmienne losowe X i Y mają rozkłady dyskretne.

Twierdzenie

Jeżeli zmienne losowe X i Y są zmiennymi losowymi typu dyskretnego, to X i Y są niezależne wtedy i tylko wtedy, gdy

$$p(x, y) = p_X(x) \cdot p_Y(y) \quad \text{dla dowolnych} \quad x \in \mathcal{X}, y \in \mathcal{Y},$$

gdzie p_X i p_Y są funkcjami rozkładów brzegowych X i Y.

Twierdzenie

Jeżeli zmienne losowe X i Y są typu ciągłego, to niezależność X i Y jest równoważna warunkowi

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

dla dowolnych $x, y \in \mathcal{R}$, dla których gęstość f jest ciągła a f_X oraz f_Y są, odpowiednio, gęstościami brzegowymi zmiennych losowych X i Y.

Mówiąc nieprecyzyjnie, zmienne losowe X i Y są niezależne, jeśli wiedza o wartości jednej zmiennej nie ma wpływu na rozkład prawdopodobieństwa drugiej.

Przykład. *Pokazać, że zmienne losowe X i Y o rozkładzie łącznym:*

	Y			
X	0	100	200	300
0	0.10	0.15	0.15	0.05
100	0.05	0.05	0.10	0.10
250	0.05	0.05	0.05	0.10

są zależne.
Znaleźć rozkład niezależnych zmiennych losowych (X_1, Y_1) mających te same rozkłady brzegowe co zmienne losowe (X, Y).
Rozwiązanie.
Ponieważ $p(0,0) = 0.10 \neq 0.45 \cdot 0.20 = p_X(0) \cdot p_Y(0)$, to X i Y są zależne.
Rozkładem łącznym zmiennych niezależnych X_1 i Y_1 jest tzw. rozkład produktowy

$$p(x,y) = p_X(x) \cdot p_Y(y) \quad \forall (x,y) \in \mathcal{R}^2.$$

Stąd

	Y_1				
X_1	0	100	200	300	$p_{X_1}(x)$
0	0.09	0.1125	0.135	0.1125	0.45
100	0.06	0.0750	0.090	0.0750	0.30
250	0.05	0.0625	0.075	0.0625	0.25
$p_{Y_1}(y)$	0.20	0.2500	0.300	0.2500	

Przykład. *Pokazać, że zmienne losowe X i Y o gęstości łącznej*

$$f(x,y) = \begin{cases} 24xy & \text{dla} \quad 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, x+y \leqslant 1 \\ 0 & \quad\quad \text{w przeciwnym przypadku.} \end{cases}$$

są zależne.

Rozwiązanie.

Jest oczywiste, że dla $(x, y) \in \{(x, y) : 0 \leqslant x \leqslant 1, 0 \leqslant y \leqslant 1, x + y > 1\}$ mamy

$$f(x, y) = 0 \neq 12x(1-x)^2 \cdot 12y(1-y)^2 = f_X(x) \cdot f_Y(y),$$

zatem X i Y są zależne.

Wniosek 1.

Jeżeli (X, Y) jest dwuwymiarową zmienną losową typu ciągłego oraz X i Y są niezależne, to obszar w którym gęstość łączna (X, Y) jest niezerowa jest postaci $\mathcal{X} \times \mathcal{Y}$, gdzie $\mathcal{X}$ jest nośnikiem X a $\mathcal{Y}$ nośnikiem Y.

Przykład. *Załóżmy, że czasy życia dwóch urządzeń są niezależne od siebie o rozkładach wykładniczych,* $X \sim Exp(\lambda_1)$ $Y \sim Exp(\lambda_2)$. *Wyznaczyć gęstość łączną zmiennej losowej* (X, Y) *oraz prawdopodobieństwo, że oba urządzenia nie zepsują się przed upływem 1500 godzin.*

Rozwiązanie.

Z niezależności X i Y otrzymujemy gęstość łączną

$$f(x,y) = f_X(x) \cdot f_Y(y) = \begin{cases} \lambda_1 e^{-\lambda_1 x} \cdot \lambda_2 e^{-\lambda_2 y} & \text{dla} \quad 0 \leqslant x, 0 \leqslant y \\ 0 & \text{w przeciwnym przypadku} \end{cases}$$

Prawdopodobieństwo zdarzenia, że oba urządzenia nie ulegną uszkodzeniu przez co najmniej 1500 godzin jest równe

$$\begin{aligned} P(X \geqslant 1500, Y \geqslant 1500) &= P(X \geqslant 1500) \cdot P(Y \geqslant 1500) = \\ &= e^{-1500(\lambda_1+\lambda_2)}. \end{aligned}$$

Na przykład, dla $\lambda_1 = 1/1000$ oraz $\lambda_2 = 1/1200$, tzn. oczekiwany czas bezawaryjnej pracy wynosi odpowiednio 1000 oraz 1200 godzin, to szukane prawdopodobieństwo wynosi

$$e^{-1500(\frac{1}{1000}+\frac{1}{1200})} = 0,2231 \cdot 0,2865 = 0,0639.$$

Rozkłady warunkowe

Związek pomiędzy dwoma zmiennymi losowymi częstokroć daje się wyjaśnić poprzez rozważenie rozkładu warunkowego jednej zmiennej pod warunkiem znajomości wartości drugiej.

Definicja.

*Jeśli X i Y są dyskretnymi zmiennymi losowymi, można zdefiniować **warunkową funkcję prawdopodobieństwa** zmiennej losowej X pod warunkiem, że $Y = y$, jako*

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}$$

dla wszystkich y takich, że $p_Y(y) > 0$.

Oczywistym jest, że

$$\sum_{x\in\mathcal{X}} p_{X|Y}(x|y) = \sum_{x\in\mathcal{X}} \frac{p(x,y)}{p_Y(y)} = \frac{1}{p_Y(y)} \sum_{x\in\mathcal{X}} p(x,y) = 1.$$

Definicja.

Jeśli X i Y są dyskretnymi zmiennymi losowymi, to **warunkową wartością oczekiwaną** *X pod warunkiem, że $Y = y$, jest*

$$E(X|Y=y) = m_X(y) = \sum_{x\in\mathcal{X}} x \cdot p_{X|Y}(x|y)$$

Przykład. *Rozważmy parę zmiennych losowych X i Y o rozkładzie łącznym podanym w tabeli:*

	Y			
X	0	100	200	300
0	0.10	0.15	0.15	0.05
100	0.05	0.05	0.10	0.10
250	0.05	0.05	0.05	0.10

Wyznaczyć rozkład warunkowy Y pod warunkiem, że $X = 100$ oraz $E(Y|X = 100)$.

Rozwiązanie.

Z definicji rozkładu warunkowego otrzymujemy

$$p_{Y|X}(y|100) = P(Y = y|X = 100) = \frac{P(X = 100, Y = y)}{P(X = 100)} = \frac{p(100, y)}{0.30}.$$

Zatem

y	0	100	200	300
$p_{Y\|X}(y\|100)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{2}{6}$

Ostatecznie mamy

$$E(Y|X=100) = 0\cdot\frac{1}{6} + 100\cdot\frac{1}{6} + 200\cdot\frac{2}{6} + 300\cdot\frac{2}{6} = \frac{1100}{6} = 183.33.$$

W przypadku zmiennych losowych typu ciągłego wygląda to nieco inaczej, ale można zdefiniować wówczas tzw. **gęstości warunkowe**

Definicja.

Jeśli X i Y są zmiennymi losowymi o gęstości łącznej $f(x, y)$ i gęstościach brzegowych $f_X(x)$ i $f_Y(y)$, to ***warunkową gęstością*** *zmiennej losowej X pod warunkiem, że $Y = y$ jest*

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)},$$

dla wszystkich $x \in R$ i y takich, że $f_Y(y) \neq 0$.

Gęstość warunkową $Y|X = x$ definiujemy analogicznie.

Oczywistym jest, że gęstość warunkowa spełnia warunki gęstości - jest nieujemna i całkuje się do 1.

$$\int_{-\infty}^{\infty} f_{X|Y}(x|y)\,dx = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f(x,y)\,dx = 1.$$

Definicja.

Jeśli X i Y są zmiennymi losowymi o gęstości łącznej $f(x, y)$, to ***waunkową wartością oczekiwaną*** *zmiennej X pod warunkiem, że $Y = y$, jest*

$$E(X|Y = y) = m_X(y) = \int_{-\infty}^{\infty} x \cdot f_{X|Y}(x|y)\, dx,$$

pod warunkiem, że całka po lewej stronie jest zbieżna bezwzględnie.

Przykład. *Jeżeli zmienne losowe X i Y są niezależne, to*

$$f_{Y|X}(y|x) = f_Y(y) \quad \textit{oraz} \quad E(Y|X = x) = E(Y).$$

Funkcje zmiennych losowych

Załóżmy, że X i Y są dwiema zmiennymi losowymi a funkcja $g(x, y)$ funkcją rzeczywistą dwóch zmiennych. Zdefiniujmy nową zmienna losową

$$Z = g(X, Y).$$

Znając rozkłady zmiennych losowych X i Y chcemy wyznaczyć rozkład prawdopodobieństwa zmiennej losowej Z.

Funkcje zmiennych losowych typu dyskretnego

Jeżeli X i Y są dwiema zmiennymi losowymi typu dyskretnego o rozkładzie łącznym $p(x, y)$, $(x, y) \in \mathcal{X} \times \mathcal{Y}$, to $Z = g(X, Y)$ jest również typu dyskretnego i dla dowolnego $z \in \mathcal{Z} = g(\mathcal{X} \times \mathcal{Y})$ mamy

Rozkład funkcji dwóch zmiennych dyskretnych

$$\forall z \in \quad P(Z = z) = \sum_{\{(x,y) \in \mathcal{X} \times \mathcal{Y}:\, g(x,y) = z\}} p(x, y).$$

Przykład. *Wyznaczyć rozkład zmiennych losowych $T = X + Y$ oraz $U = max(X, Y)$, gdzie rozkład (X, Y) podany jest w tabeli poniżej.*

	Y			
X	*0*	*1*	*2*	*3*
0	*0.1*	*0.1*	*0.2*	*0.1*
1	*0.1*	*0.1*	*0.1*	*0.0*
2	*0.1*	*0.0*	*0.0*	*0.1*

Rozwiązanie.
Nośnikiem zmiennej losowej T jest $\mathcal{T} = \{0, 1, 2, 3, 4, 5\}$.

$$\{(x, y) : x + y = 0\} = \{(0, 0)\}, \ \text{zatem} \ P(T = 0) = p(0, 0) = 0.1.$$

Podobnie otrzymujemy, że

$$
\begin{aligned}
\{(x,y): x+y=1\} &= \{(0,1),(1,0)\} \\
& \quad P(T=1) = p(0,1)+p(1,0) = 0.2, \\
\{(x,y): x+y=2\} &= \{(0,2),(2,0),(1,1)\} \\
& \quad P(T=2) = p(0,2)+p(2,0)+p(1,1) = 0.4, \\
\{(x,y): x+y=3\} &= \{(1,2),(2,1),(0,3)\} \\
& \quad P(T=2) = p(1,2)+p(2,1)+p(0,3) = 0.2, \\
\{(x,y): x+y=4\} &= \{(2,2),(1,3)\} \\
& \quad P(T=2) = p(2,2)+p(1,3) = 0.0, \\
\{(x,y): x+y=5\} &= \{(2,3)\} \\
& \quad P(T=5) = p(2,3) = 0.1,
\end{aligned}
$$

Ostatecznie

t	0	1	2	3	5
$p_T(t)$	0.1	0.2	0.4	0.2	0.1

Dla zmiennej losowej U nośnikiem jest $\mathcal{U} = \{0, 1, 2, 3\}$.

$$
\begin{aligned}
\{(x,y) : \max(x,y) = 0\} &= \{(0,0)\} \\
\{(x,y) : \max(x,y) = 1\} &= \{(0,1),(1,0),(1,1)\} \\
\{(x,y) : \max(x,y) = 2\} &= \{(0,2),(2,0),(1,2),(2,1),(2,2)\} \\
\{(x,y) : \max(x,y) = 3\} &= \{(0,3),(1,3),(2,3)\}
\end{aligned}
$$

Funkcja prawdopodobieństwa zmiennej U

u	0	1	2	3
$p_U(u)$	0.1	0.3	0.4	0.2

Wartość oczekiwana, kowariancja, korelacja

Dla zmiennej losowej jednowymiarowej X zdefiniowaliśmy wartość oczekiwaną $E[h(X)]$ dowolnej funkcji $h(X)$ od tej zmiennej losowej jako średnią ważoną wartości tej funkcji, z wagami będącymi funkcją prawdopodobieństwa $p(x)$ lub gęstością prawdopodobieństwa $f(x)$ zmiennej X.

Podobną definicję podamy dla funkcji dwuwymiarowej $h : R^2 \rightarrow R$ od zmiennej losowej dwuwymiarowej (X, Y).

Niech X i Y będą zmiennymi losowymi o rozkładzie łącznym zdefiniowanym przez funkcję prawdopodobieństwa $p(x, y)$ lub funkcję gęstości $f(x, y)$ w zależności czy X i Y są typu dyskretnego czy typu ciągłego.

Definicja.

Wartością oczekiwaną funkcji $h(X, Y)$, oznaczoną przez $E[h(X, Y)]$ nazywamy

$$E[h(X,Y)] = \begin{cases} \sum_x \sum_y h(x,y) \cdot p(x,y) & X, Y \text{ typu dyskret.} \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x,y) \cdot f(x,y) dxdy & X, Y \text{ typu ciągłego} \end{cases}$$

pod warunkiem, ze szereg (całka) po prawej stronie jest zbieżny (zbieżna) absolutnie.

Rozważmy zmienną losową

$$h(X, Y) = X,$$

w przypadku, gdy (X, Y) jest typu ciągłego o gęstości łącznej $f(x, y)$, wówczas

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \cdot f(x, y)dxdy = \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x, y)dy \right) dx = \\ &= \int_{-\infty}^{\infty} x \cdot f_X(x)dx, \end{aligned}$$

gdzie $f_X(x)$ jest gęstością brzegową zmiennej X.
Wynika stąd, że wartość oczekiwana rozkładu brzegowego X jest taka sama jak w przypadku jednowymiarowym.

Wykorzystując definicję wartości oczekiwanej funkcji zmiennych losowych dwuwymiarowych wyznaczymy wartość oczekiwaną oraz wariancję kombinacji liniowej dwóch zmiennych losowych.

Twierdzenie.

Dla dowolnych zmiennych losowych (X, Y) oraz dowolnych stałych $a, b \in R$

$$E(aX + bY) = aE[X] + bE[Y].$$

Dowód.

Załóżmy, że (X, Y) jest typu ciągłego o gęstości $f(x, y)$ (w przypadku dyskretnym dowód jest identyczny, jedynie całkę należy zastąpić sumą), wówczas

$$\begin{aligned}
E[aX + bY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(ax + by)\cdot f(x, y)dxdy = \\
&= a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot f(x, y)dxdy + b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\cdot f(x, y)dxdy \\
&= aE[X] + bE[Y].
\end{aligned}$$

Twierdzenie.
Dla dowolnych zmiennych losowych (X, Y) oraz dowolnych stałych $a, b \in R$

$$V(aX + bY) = a^2 V[X] + b^2 V[Y] + 2ab \cdot E\{(X - E[X])(Y - E[Y])\}.$$

Dowód.
Z definicji wariancji otrzymujemy

$$\begin{aligned}
V(aX + bY) &= E\{[(aX + bY) - E(aX + bY)]^2\} = \\
&= E\{[a(X - E[X]) + b(Y - E[Y])]^2\} = \\
&= E\{a^2(X - E[X])^2 + b^2(Y - E[Y])^2 + \\
&\quad +2ab(X - E[X])(Y - E[Y])\} = \\
&= a^2 V[X] + b^2 V[Y] + \\
&\quad +2ab \cdot E\{(X - E[X])(Y - E[Y])\}.
\end{aligned}$$

Wariancja kombinacji liniowej dwóch zmiennych losowych zależy od obu wariancji oraz dodatkowo od wielkości

$$E\{(X - E[X])(Y - E[Y])\},$$

która mówi o wzajemnej zależności pomiędzy zmiennymi X i Y.

Definicja.

Kowariancją *zmiennych losowych X i Y nazywamy wyrażenie*

$$Cov(X, Y) = E\{(X - E[X])(Y - E[Y])\}.$$

W ogólności, dodatnia wartość $Cov(X, Y)$ oznacza, że zmienna losowa Y raczej rośnie wraz ze wzrostem wartości X, natomiast w przypadku, gdy kowariancja jest ujemna, raczej maleje.

Można powiedzieć, że kowariancja jest miarą zależności pomiędzy X i Y. Jednakże jej wartość mocno zależy od bezwzględnej wielkości X i Y.

Aby usunąć tę wadę, zamiast kowariancji zmiennych X i Y rozważa się kowariancję zmiennych standaryzowanych X^* i Y^*, która mierzy siłę wzajemnej korelacji pomiędzy X i Y.

Definicja.

Współczynnikiem korelacji *zmiennych losowych X i Y, oznaczonym jako $Corr(X, Y)$, $\rho_{X,Y}$ lub ρ, nazywamy*

$$\rho_{X,Y} = Cov\left(\frac{X - E[X]}{\sigma_X}, \frac{Y - E[Y]}{\sigma_Y}\right) = \frac{Cov(X, Y)}{\sigma_X \cdot \sigma_Y}.$$

Twierdzenie.
Jeżeli (X, Y) są zmiennymi losowymi niezależnymi, dla których istnieje $E(X \cdot Y)$, to

$$E(X \cdot Y) = E[X] \cdot E[Y].$$

Dowód.
Dowód przeprowadzimy w przypadku, gdy (X, Y) jest typu ciągłego o gęstości $f(x, y) = f_X(x)f_Y(y)$ (w przypadku dyskretnym jest podobny)

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot f(x, y)dxdy = \\ &= \left(\int_{-\infty}^{\infty} x \cdot f_X(x)dx\right) \cdot \left(\int_{-\infty}^{\infty} y \cdot f_Y(y)dy\right) \\ &= E[X] \cdot E[Y]. \end{aligned}$$

Definicja.

Mówimy, że zmienne losowe X i Y są ***nieskorelowane*** *jeżeli*

$$Cov(X, Y) = 0.$$

Wniosek

Jeżeli zmienne losowe X i Y są niezależne oraz istnieje ich kowariancja, to są one nieskorelowane.

Twierdzenie odwrotne nie jest prawdziwe.

Własności współczynnika korelacji

Dla dowolnych zmiennych losowych X i Y, dla których istnieje współczynnik korelacji ρ, mamy

1. $|\rho_{XY}| \leqslant 1$.
2. Jeżeli X i Y są niezależne, to $\rho_{XY} = 0$.
3. $|\rho_{XY}| = 1$ wtedy i tylko wtedy, gdy istnieją takie $a \neq 0$ oraz b, że $P(Y = aX + b) = 1$.
4. $\forall a \neq 0, c \neq 0$ i $\forall b, d$ mamy

$$|Corr(X, Y)| = |Corr(aX + b, cY + d)|.$$