

ROZKŁAD PRAWDOPODOBIEŃSTWA WIELU ZMIENNYCH LOSOWYCH

W przypadku gdy mamy do czynienia z n zmiennymi losowymi możemy prawdopodobieństwo, iż przyjmą one wartości $x_1, x_2, \dots, x_n$ opisać wielowymiarową funkcją rozkładu gęstości prawdopodobieństwa $f(x_1, x_2, \dots, x_n)$.

W szczególnym przypadku, gdy zmienne te są niezależne, funkcję $f(x_1, x_2, \dots, x_n)$ można przedstawić jako iloczyn funkcji gęstości prawdopodobieństwa poszczególnych zmiennych:

$$f(x_1, x_2, \dots, x_n) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_n)$$

Warunek normalizacyjny funkcji $f(x_1, x_2, \dots, x_n)$, wynikający z definicji prawdopodobieństwa ma dla wielowymiarowej funkcji rozkładu postać:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n) dx_1 \cdot dx_2 \cdot \dots \cdot dx_n = 1$$

Wartość oczekiwaną zmiennej x_i definiujemy wzorem:

$$E(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i \cdot f(x_1, x_2, \dots, x_n) dx_1 \cdot dx_2 \cdot \dots \cdot dx_n$$

Możemy także wprowadzić pojęcie wartości oczekiwanej dowolnej funkcji zmiennych losowych $x_1, x_2, \dots, x_n$. Jeśli funkcję tą oznaczmy jako $y(x_1, x_2, \dots, x_n)$ to możemy napisać:

$$E(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} y(x_1, x_2, \dots, x_n) \cdot f(x_1, x_2, \dots, x_n) dx_1 \cdot dx_2 \cdot \dots \cdot dx_n$$

Przykładami takich funkcji są wzory na estymatory wartości oczekiwanej (wartość średnia) czy dyspersji (odchylenie standardowe).

WARIANCJA, KOWARIANCJA

Podobnie jak wartość oczekiwaną zmiennej losowej x_i możemy zdefiniować jej wariancję:

$$V(x_i) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [x_i - E(x_i)]^2 \cdot f(x_1, x_2, \dots, x_n) dx_1 \cdot dx_2 \cdot \dots \cdot dx_n$$

Dla rozkładów prawdopodobieństwa wielu zmiennych losowych możemy jednak pojęcie wariancji rozszerzyć na przypadek dwóch dowolnych zmiennych losowych wprowadzając tym samym pojęcie kowariancji.

Kowariancja zmiennych losowych x_i i x_j jest zdefiniowana następująco:

$$\text{cov}(x_i, x_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} [x_i - E(x_i)] \cdot [x_j - E(x_j)] \cdot f(x_1, x_2, \dots, x_n) dx_1 \cdot dx_2 \cdot \dots \cdot dx_n$$

Kowariancja informuje nas o tym, czy zmienne x_i i x_j są ze sobą powiązane. Jeżeli zmienne te są niezależne to ich kowariancja = 0.

Łatwo zauważyć że: $\text{cov}(x_i, x_i) = V(x_i)$ **oraz** $\text{cov}(x_i, x_j) = \text{cov}(x_j, x_i)$

Zbiór wartości $\text{cov}(x_i, x_j)$ dla $i, j = 1, 2, \dots, n$ nazywa się macierzą kowariancji.

Ponieważ $\text{cov}(x_i, x_i) = V(x_i)$ macierz kowariancji jest symetryczna.

$$\begin{bmatrix} V(x_1) & \text{cov}(x_1, x_2) & \text{cov}(x_1, x_3) & \dots & \text{cov}(x_1, x_n) \\ \text{cov}(x_2, x_1) & V(x_2) & \text{cov}(x_2, x_3) & \dots & \text{cov}(x_2, x_n) \\ \text{cov}(x_3, x_1) & \text{cov}(x_3, x_2) & V(x_3) & \dots & \text{cov}(x_3, x_n) \\ \dots & \dots & \dots & \dots & \dots \\ \text{cov}(x_n, x_1) & \text{cov}(x_n, x_2) & \text{cov}(x_n, x_3) & \dots & V(x_n) \end{bmatrix}$$

WSPÓŁCZYNNIK KORELACJI

Oprócz kowariancji definiuje się także inne narzędzie przydatne do określenia powiązania pomiędzy zmiennymi losowymi. Jest nim **współczynnik korelacji** zdefiniowany jako:

$$r_{ij} = \frac{\text{cov}(x_i, x_j)}{\sqrt{V(x_i) \cdot V(x_j)}}$$

Współczynnik korelacji przyjmuje wartości z zakresu od -1 do 1.

Jeżeli $r_{ij} > 0$ oznacza to, że wzrostowi wartości x_i towarzyszy wzrost wartości x_j - określamy je wtedy jako dodatnio skorelowane.

Jeżeli $r_{ij} < 0$ to wzrostowi wartości x_i towarzyszy spadek wartości x_j - zmienne te określa się jako skorelowane ujemnie.

W przypadku gdy $r_{ij} = 0$ mówimy, że zmienne x_i oraz x_j są nieskorelowane.

W szczególności $r_{ij} = 0$ dla zmiennych niezależnych, jednakże fakt zerowej wartości współczynnika korelacji nie jest równoważny niezależności zmiennych

(tzn. może zdarzyć się tak, że współczynnik korelacji jest równy zero, a zmienne są od siebie zależne).

PRAWO PROPAGACJI NIEPEWNOŚCI POMIAROWYCH

Oprócz wielkości mierzonych bezpośrednio w praktyce pomiarowej często mamy do czynienia z wielkościami, które wyznaczamy pośrednio. Przeważnie mierzymy wówczas kilka innych wielkości $x_1, x_2, \dots, x_n$ po czym na podstawie otrzymanych wyników wyliczamy ze wzoru wartość interesującej nas wielkości $y = y(x_1, x_2, \dots, x_n)$. Powstaje zatem pytanie jak na podstawie wartości niepewności pomiarów - $u(x_1), u(x_2), \dots, u(x_n)$ wyznaczyć niepewność $u(y)$.

Ponieważ w omawianym statystycznym modelu opisu niepewności jako niepewność przyjmujemy oszacowanie dyspersji możliwych wartości wielkości mierzonej, również w tym przypadku interesować nas będzie dyspersja oraz wariancja zmiennej losowej y . Okazuje się, że rozważania analityczne możemy przeprowadzić tylko wówczas, gdy zależność $y = y(x_1, x_2, \dots, x_n)$ jest liniowa lub gdy możemy ją traktować jako liniową w zakresie zmian wartości zmiennych losowych $x_1, x_2, \dots, x_n$ wynikających z rozrzutu każdej z nich (tzn. gdy niepewności pomiarów - $u(x_1), u(x_2), \dots, u(x_n)$ są stosunkowo małe).

W celu uproszczenia przekształceń rozważmy nieliniową funkcję dwóch zmiennych losowych x_1, x_2 , którą oznaczmy $y = y(x_1, x_2)$.

Przedstawmy funkcję $y(x_1, x_2)$ w postaci rozwinięcia w szereg wokół wartości oczekiwanych zmiennych $x_1, x_2 - \mu_1, \mu_2$:

$$y(x_1, x_2) = y(\mu_1, \mu_2) + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) + \dots$$

Jeżeli przyjmiemy, że rozrzut zmiennych x_1, x_2 jest mały, wówczas możemy w powyższym rozwinięciu pozostawić tylko wyrazy liniowe.

Wyznamy teraz wariancję $V(y)$. W tym celu skorzystamy z zależności:

$$V(y) = E(y^2) - (E(y))^2$$

$$\begin{aligned} E(y) &= E \left\{ y(\mu_1, \mu_2) + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) \right\} = \\ &= E(y(\mu_1, \mu_2)) + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} E(x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} E(x_2 - \mu_2) = \\ &= y(\mu_1, \mu_2) \end{aligned}$$

Zatem:

$$(E(y))^2 = (y(\mu_1, \mu_2))^2$$

W celu wyznaczenia $E(y^2)$ wyliczmy najpierw y^2 .

$$\begin{aligned}
y^2 &= \left\{ y(\mu_1, \mu_2) + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) \right\}^2 = \\
&= y(\mu_1, \mu_2)^2 + 2y(\mu_1, \mu_2) \left\{ \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) \right\} + \\
&+ \left\{ \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) \right\}^2 = \\
&= y(\mu_1, \mu_2)^2 + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} (x_1 - \mu_1) + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_2 - \mu_2) + \\
&+ \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 (x_1 - \mu_1)^2 + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} (x_1 - \mu_1)(x_2 - \mu_2) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 (x_2 - \mu_2)^2
\end{aligned}$$

Zatem:

$$\begin{aligned}
E(y^2) &= E(y(\mu_1, \mu_2)^2) + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} E(x_1 - \mu_1) + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} E(x_2 - \mu_2) + \\
&+ \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 E[(x_1 - \mu_1)^2] + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} E[(x_1 - \mu_1)(x_2 - \mu_2)] + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 E[(x_2 - \mu_2)^2]
\end{aligned}$$

$$E(y^2) = E(y(\mu_1, \mu_2)^2) + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} E(x_1 - \mu_1) + 2y(\mu_1, \mu_2) \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} E(x_2 - \mu_2) +$$

$$+ \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 E[(x_1 - \mu_1)^2] + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} E[(x_1 - \mu_1)(x_2 - \mu_2)] + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 E[(x_2 - \mu_2)^2]$$

Ponieważ:

$$E(y(\mu_1, \mu_2))^2 = y(\mu_1, \mu_2)^2$$

$$E(x_1 - \mu_1) = 0$$

$$E(x_2 - \mu_2) = 0$$

$$E[(x_1 - \mu_1)^2] = V(x_1)$$

$$E[(x_2 - \mu_2)^2] = V(x_2)$$

$$E[(x_1 - \mu_1)(x_2 - \mu_2)] = \text{cov}(x_1, x_2)$$

zatem:

$$E(y^2) = y(\mu_1, \mu_2)^2 + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 V(x_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 V(x_2) + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} \text{cov}(x_1, x_2)$$

oraz

$$V(y) = E(y^2) - (E(y))^2 =$$

$$= y(\mu_1, \mu_2)^2 + \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 V(x_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 V(x_2) + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} \text{cov}(x_1, x_2) - y(\mu_1, \mu_2)^2 =$$

$$= \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 V(x_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 V(x_2) + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} \text{cov}(x_1, x_2)$$

Otrzymana zależność to uogólniona postać **prawa propagacji wariancji**. -

Jeżeli $y=y(x_1, x_2)$ to

$$V(y) = \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 V(x_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 V(x_2) + 2 \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1} \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2} \text{cov}(x_1, x_2)$$

Zatem jeżeli przyjmiemy, że zmierzaliśmy bezpośrednio dwie wielkości i otrzymaliśmy w wyniku pomiarów wartości x_{1sr} i x_{2sr} z ich niepewnościami oraz wyliczamy wartość y to **prawo propagacji niepewności** możemy sformułować jako:

$$u(y) = \sqrt{\left(\frac{\partial y}{\partial x_1} \right)_{x_{1sr}}^2 u(x_{1sr})^2 + \left(\frac{\partial y}{\partial x_2} \right)_{x_{2sr}}^2 u(x_{2sr})^2 + 2 \left(\frac{\partial y}{\partial x_1} \right)_{x_{1sr}} \left(\frac{\partial y}{\partial x_2} \right)_{x_{2sr}} \text{cov}(x_1, x_2)}$$

Jeżeli pomiary wielkości x_1, x_2 będą niezależne (z takim przypadkiem mamy najczęściej do czynienia) to wzory powyższe przyjmują postać:

$$V(y) = \left(\frac{\partial y}{\partial x_1} \right)_{\mu_1}^2 V(x_1) + \left(\frac{\partial y}{\partial x_2} \right)_{\mu_2}^2 V(x_2)$$

$$u(y) = \sqrt{\left(\frac{\partial y}{\partial x_1} \right)_{x_{1sr}}^2 u(x_{1sr})^2 + \left(\frac{\partial y}{\partial x_2} \right)_{x_{2sr}}^2 u(x_{2sr})^2}$$