

Definicja. Zmienną losową X zdefiniowaną na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$ nazywamy funkcję $X : \Omega \rightarrow \mathbb{R}$ taką, że $X^{-1}(B) \in \mathcal{F}$ dla każdego zbioru borelowskiego $B \subseteq \mathbb{R}$. Innymi słowy, zmienna losowa to funkcja mierzalna z przestrzeni $(\Omega, \mathcal{F}, \mathbb{P})$ do zbioru liczb rzeczywistych $\mathbb{R}$.

Uwaga 1. Ze zmienną losową X związana jest nowa przestrzeń probabilistyczna $(\mathbb{R}, \mathcal{B}, \mathbb{P}_X)$, gdzie $\mathcal{B}$ jest σ -algebrą zbiorów borelowskich, a

$$\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B)) = \mathbb{P}(X \in B).$$

Jeśli interpretujemy X jako wygraną, której wynik zależy od pewnego eksperymentu losowego, to $\mathbb{P}(X \in B)$ mówi o prawdopodobieństwie, że wygrana wypadnie do zbioru B .

Definicja. Dystrybuantą zmiennej losowej X nazywamy funkcję $F_X : \mathbb{R} \rightarrow [0, 1]$ zdefiniowaną wzorem

$$F_X(x) = \mathbb{P}(X \leq x).$$

Dystrybuanta jednoznacznie wyznacza miarę $\mathbb{P}_X$ generowaną przez X na zbiorze $\mathbb{R}$.

Twierdzenie 1 (WŁASNOŚCI DYSTRYBUANTY). Jeśli $F : \mathbb{R} \rightarrow [0, 1]$ jest dystrybuantą pewnej zmiennej losowej, to

- (i) F jest niemalejąca,
- (ii) F jest prawostronnie ciągła,
- (iii) $\lim_{x \rightarrow -\infty} F(x) = 0$ i $\lim_{x \rightarrow \infty} F(x) = 1$.

Co więcej, każda funkcja spełniająca warunki (i)–(iii) jest dystrybuantą pewnej zmiennej losowej.

Definicja. Jeśli miara $\mathbb{P}_X$ jest **czysto atomowa**, tzn. istnieje przeliczalny zbiór **atomów** $A \subseteq \mathbb{R}$, dla którego

$$\sum_{a \in A} \mathbb{P}_X(a) = \sum_{a \in A} \mathbb{P}(X = a) = 1,$$

to X nazywamy **zmienną losową dyskretną**. Dystrybuanta takiej zmiennej losowej jest funkcją schodkową.

Definicja. Jeśli miara $\mathbb{P}_X$ jest **absolutnie ciągła** względem miary Lebesgue’a λ (tzn. z faktu $\lambda(B) = 0$ wynika, że $\mathbb{P}(B) = 0$), to z twierdzenia Radona-Nikodyma istnieje **funkcja gęstości** $f : \mathbb{R} \rightarrow [0, \infty]$ taka, że dla każdego zbioru mierzalnego B

$$\mathbb{P}(X \in B) = \int_B f(x) dx.$$

Mówimy wtedy, że X jest **zmienną losową ciągłą**. Dla takich zmiennych losowych

$$F(x) = \int_{-\infty}^x f(t) dt$$

i

$$f(x) = F'(x).$$

Należy pamiętać o tym, że o ile dystrybuanta $F(x)$ jest określona jednoznacznie, gęstość $f(x)$, która występuje pod całką, jest określona z dokładnością do zbiorów miary (Lebesgue’a) zero.

Definicja. Jeśli miara $\mathbb{P}_X$ jest bezatomowa, ale jest singularna, tzn. istnieje zbiór B , którego miara Lebesgue’a jest równa zero, dla którego $\mathbb{P}(B) = 1$, to X jest **zmienną losową singularną**. Dystrybuanta takiej zmiennej jest funkcją ciągłą, której pochodna jest równa zero wszędzie, poza zbiorem miary zero.

Twierdzenie 2 (TWIERDZENIE O ROZKŁADZIE LEBESGUE'A). Każdą miarę probabilistyczną $\mathbb{P}$ na zbiorze $\mathbb{R}$ można przedstawić w postaci

$$\mathbb{P} = a\mathbb{P}^d + b\mathbb{P}^{ac} + c\mathbb{P}^{sing},$$

gdzie $a, b, c \geq 0$ to stałe takie, że $a + b + c = 1$, $\mathbb{P}^d$ jest miarą czysto atomową, $\mathbb{P}^{ac}$ miarą absolutnie ciągłą, a $\mathbb{P}^{sing}$ miarą singularną.

Definicja. Jeśli X jest zmienną losową określoną na przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P})$ i

$$\int_{\Omega} |X(\omega)| dP(\omega) < \infty,$$

to liczbę

$$(1) \quad \mathbb{E}X = \int_{\Omega} |X(\omega)| dP(\omega) = \int_{-\infty}^{\infty} x dF_X$$

nazywamy **wartością oczekiwaną** zmiennej losowej X . Dla zmiennych losowych dyskretnych powyższa definicja przyjmuje postać

$$\mathbb{E}X = \sum_{a \in A} a \mathbb{P}(X = a),$$

natomiast dla zmiennych losowych ciągłych mamy

$$\mathbb{E}X = \int_{-\infty}^{\infty} t f(t) dt.$$

Twierdzenie 3 (PRAWO LENIWEGO STATYSTYKA). Jeśli X jest zmienną losową, a $g : \mathbb{R} \rightarrow \mathbb{R}$ jest funkcją mierzalną, to

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(x) dF_X.$$

W szczególności, dla zmiennych losowych dyskretnych

$$\mathbb{E}g(X) = \sum_{a \in A} g(a) \mathbb{P}(X = a),$$

a dla zmiennych losowych ciągłych

$$\mathbb{E}g(X) = \int_{-\infty}^{\infty} g(t) f(t) dt.$$

Definicja. Wariancję zmiennej losowej X definiujemy jako

$$\text{Var}X = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}(X^2) - (\mathbb{E}X)^2.$$

Nieco bardziej ogólnie, dla $k \geq 1$, **k -ty moment** X jest zdefiniowany jako $\mathbb{E}X^k$, **k -ty moment silniowy** X to $\mathbb{E}X(X-1)\cdots(X-k+1)$, a **k -ty moment centralny** określamy jako $\mathbb{E}(X - \mathbb{E}X)^k$. Zatem, np. wariancja to drugi moment centralny. Oczywiście istnieją zmienne losowe nie mające k -tego momentu, ale jeśli k -ty moment (dowolnego typu) istnieje, to istnieją również momenty niższego rzędu (wynika to z wniosku z nierówności Jensena podanego poniżej).

POŻYTECZNE NIERÓWNOŚCI

NIERÓWNOŚĆ MARKOWA Jeśli X jest zmienną losową i $a > 0$, to

$$\mathbb{P}(|X| \geq a) \leq \frac{\mathbb{E}|X|}{a}.$$

NIERÓWNOŚĆ CZEBYSZEWA Jeśli Y jest zmienną losową i $b > 0$, to

$$\mathbb{P}(|Y - \mathbb{E}Y| \geq b) \leq \frac{\text{Var}Y}{b^2}.$$

Dowód: Wystarczy podstawić $X = (Y - \mathbb{E}Y)^2$ i $a = b^2$ do nierówności Markowa. $\square$