Metody Probabilistyczne i Statystyka - Wykład 4.

Charakterystyki liczbowe jednowymiarowych zmiennych losowych

17 marca 2025

Przykład 1.

Rzucamy 1 raz kostką sześcienną. Jaka będzie średnia liczba oczek?

Wartość oczekiwana

Definicja

Niech X będzie jednowymiarową zmienną losową o rozkładzie dyskretnym. Jeśli szereg

$$\sum_{x_i \in S_X} |x_i| \cdot P(X = x_i)$$

jest zbieżny, to **wartością oczekiwaną** *zmiennej losowej X nazywamy liczbę*

$$EX = \sum_{x_i \in S_X} x_i \cdot P(X = x_i).$$

Jeśli szereg $\sum_{x_i \in S_X} |x_i| \cdot P(X = x_i)$ nie jest zbieżny to mówimy, że EX nie istnieje.

Wartość oczekiwana

Interpretacja wartości oczekiwanej w przypadku dyskretnym:

Wartość oczekiwana - środek ciężkości układu składającego się z obiektów rozmieszczonych w punktach $x \in S_X$ o masach równych $P(X = x)$.

Przykład 2.

Rzucamy 1 raz sześcienną prawidłową kostką. Jeśli wypadnie "6", wygrywamy 50 zł. Jeśli wypadnie nieparzysta liczba oczek, przegrywamy 60 zł. W pozostałych przypadkach wygrywamy 10 zł. Niech X oznacza wygraną. Wtedy:

$$P(X = -60) = \frac{1}{2}, \; P(X = 10) = \frac{1}{3}, \; P(X = 50) = \frac{1}{6}.$$

Wyznaczyć średnią wygraną.

Wartość oczekiwana

Definicja

Niech X będzie jednowymiarową zmienną losową o rozkładzie ciągłym. Jeśli całka

$$\int\limits_{-\infty}^{+\infty} |x| \cdot f_X(x) dx$$

jest zbieżna, to **wartością oczekiwaną** *zmiennej losowej X nazywamy liczbę*

$$EX = \int\limits_{-\infty}^{+\infty} x \cdot f_X(x) dx.$$

Jeśli całka $\int\limits_{-\infty}^{+\infty} |x| \cdot f_X(x) dx$ *nie jest zbieżna to mówimy, że EX nie istnieje.*

Wartość oczekiwana

Interpretacja wartości oczekiwanej w przypadku ciągłym:

Wartość oczekiwana - środek ciężkości masy prawdopodobieństwa opisanej przez funkcję gęstości.

Przykład 3.

Niech T oznacza czas oczekiwania na autobus mający przyjechać w ciągu godziny. Ile średnio czasu będziemy czekać?

Wartość oczekiwana

Wartość oczekiwana dla próby

- Przypuśćmy, że X oznacza pewną cechę populacji.
- Z tej populacji losujemy próbę n-elementową otrzymując wyniki $x_1, \dots, x_n$, gdzie x_i oznacza wartość cechy X dla i-tego elementu.
- Wtedy **wartość średnią dla wylosowanej próby** oznaczamy symbolem $\bar{x}$ i

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i.$$

Uwagi

- Średnia z próby jest dobrym oszacowaniem wartości oczekiwanej dla całej populacji.
  - EX – wartość średnia cechy X mierzona dla całej populacji
  - $\bar{x}$ – wartość średnia cechy X mierzona dla wylosowanej próby

Własności wartości oczekiwanej

Twierdzenie

Niech X i Y będą zmiennymi losowymi, dla których istnieją EX i EY. Wtedy:

1. $E(b) = b$ *dla każdego* $b \in \mathbb{R}$
2. $E(a \cdot X) = a \cdot EX$ *dla każdego* $a \in \mathbb{R}$
3. $E(X + Y) = EX + EY$
4. *Jeśli* $X \geqslant 0$, *to* $EX \geqslant 0$

Wartość oczekiwana funkcji zmiennej losowej

Twierdzenie

Niech $g : \mathbb{R} \to \mathbb{R}$ będzie funkcją taką, że $g(X)$ jest zmienną losową jednowymiarową.

- *Jeśli X ma rozkład dyskretny, to*

$$E\left(g(X)\right) = \sum_{x_i \in S_X} g(x_i) \cdot P(X = x_i).$$

- *Jeśli X ma rozkład ciągły, to*

$$E\left(g(X)\right) = \int_{-\infty}^{+\infty} g(x) \cdot f_X(x) dx.$$

Przykład 4.

Zmienna losowa X ma rozkład ciągły o gęstości

$$f_X(x) = \frac{1}{2} \cdot 1_{[0;2]}(x).$$

Niech $Y = g(X) = \max(1, X)$. Wyznaczyć EY.

Wariancja i odchylenie standardowe

Przykład 5.

Rozważmy inwestycję, z której zysk jest zmienną losową X o rozkładzie

1. $$P(X = 1\,000) = P(X = -800) = \frac{1}{2}$$
2. $$P(X = 101) = P(X = 99) = \frac{1}{2}$$

Wariancja i odchylenie standardowe

Definicja

- *Jeśli istnieje EX^2, to liczbę*

$$Var(X) = E(X - EX)^2,$$

nazywamy **wariancją** *zmiennej losowej X.*

- *Wtedy liczbę $\sigma_X = \sqrt{Var(X)}$ nazywamy* **odchyleniem standardowym** *zmiennej losowej X.*

Uwagi

- Odchylenie standardowe przyjmuje zazwyczaj wartości bliskie $E|X - EX|$.
- Wariancja zmiennej losowej X często oznaczana jest symbolem VX.

Wariancja i odchylenie standardowe

Uwaga

Wariancja i odchylenie standardowe są miarami rozrzutu wartości zmiennej losowej w stosunku do wartości średniej.

Twierdzenie

Jeśli istnieje EX^2, to

$$Var(X) = EX^2 - (EX)^2.$$

Własności wariancji

Twierdzenie

Jeśli X jest zmienną losową, dla której istnieje EX^2, to:

1. $Var(X) \geqslant 0$
2. $Var(X) = 0 \Leftrightarrow X$ *ma rozkład jednopunktowy, to znaczy* $P(X = a) = 1$ *dla pewnego* $a \in \mathbb{R}$
3. $Var(aX + b) = a^2 \cdot Var(X)$ *dla wszystkich* $a, b \in \mathbb{R}$
4. *W ogólnym przypadku* $Var(X + Y) \neq Var(X) + Var(Y)$

Wariancja i odchylenie standardowe

Uwaga

- **Wariancja dla wylosowanej próby** oznaczana jest symbolem s^2 i

$$s^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

- Wtedy $s = \sqrt{s^2}$ - **odchylenie standardowe z próby**.

Uwagi

- s^2 i s są dobrymi oszacowaniami wariancji i odchylenia standardowego dla całej populacji.
  - $Var(X) -$ wariancja cechy X mierzona dla całej populacji
  - $s^2 -$ wariancja cechy X mierzona dla wylosowanej próby

Nierówność Czebyszewa

Twierdzenie

Dla każdego $\varepsilon > 0$

$$P\left(|X - EX| \geqslant \varepsilon\right) \leqslant \frac{\sigma_X^2}{\varepsilon^2}.$$

Jeśli $\varepsilon = k \cdot \sigma_X$, *to*

$$P\left(|X - EX| < k \cdot \sigma_X\right) \geqslant 1 - \frac{1}{k^2}.$$

Uwaga

Przedział $[\bar{x} - 2s; \bar{x} + 2s]$ nazywamy **przedziałem wartości typowych** cechy X (obserwacje znajdujące się poza tym przedziałem to **obserwacje odstające**)