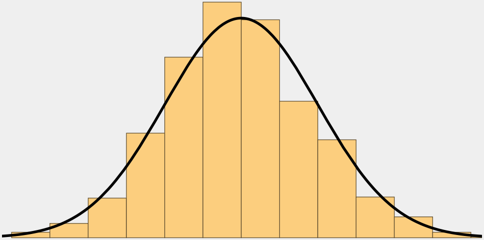
Probabilistyka

Wykład piąty: Porozmawiajmy o Centralnym

Grzegorz Siudem

Wydział Fizyki

wykład zdalny 2020



W poprzednim odcinku...

Na poprzednim wykładzie opowiadałem o

- Prawdopodobieństwie warunkowym.
- wzorze Bayesa i prawdopodobieństwie całkowitym.
- wnioskowaniu Bayesowskim.

Testy diagnostyczne

Definicje (proszę mi przypomnieć o rysunku!)

- Czułość testu to odsetek chorych z pozytywnym wynikiem.
- Swoistość testu to odsetek zdrowych z wynikiem ujemnym.

Zadanie - próba wysiłkowa (czarny J-Sz, str. 59)

Próba wysiłkowa, której czułość wynosi 65%, a swoistość 85% jest powszechnie używana w celu wykrycia choroby wieńcowej. Załóżmy, że (w populacji) jest 10% chorych na chorobę wieńcową. Obliczmy

- p-stwo, że próba prowadzi do prawidłowej diagnozy.

Zadanie - próba wysiłkowa (czarny J-Sz, str. 59)

Próba wysiłkowa, której czułość wynosi 65%, a swoistość 85% jest powszechnie używana w celu wykrycia choroby wieńcowej. Załóżmy, że (w populacji) jest 10% chorych na chorobę wieńcową. Obliczmy

- p-stwo, że próba prowadzi do prawidłowej diagnozy.
- p-stwo, że pacjent z wynikiem dodatnim jest chory.

Zadanie - próba wysiłkowa (czarny J-Sz, str. 59)

Próba wysiłkowa, której czułość wynosi 65%, a swoistość 85% jest powszechnie używana w celu wykrycia choroby wieńcowej. Załóżmy, że (w populacji) jest 10% chorych na chorobę wieńcową. Obliczmy

- p-stwo, że próba prowadzi do prawidłowej diagnozy.
- p-stwo, że pacjent z wynikiem dodatnim jest chory.
- p-stwo, że pacjent z wynikiem ujemnym jest zdrowy.

Twierdzenia graniczne

Twierdzenie

Jeśli $n \to \infty$, $p_n \to 0$ oraz $np_n \to \lambda > 0$ to

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} \xrightarrow{n \to \infty} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Twierdzenie

Jeśli $n \to \infty$, $p_n \to 0$ oraz $np_n \to \lambda > 0$ to

$$\binom{n}{k} p_n^k (1 - p_n)^{n-k} \xrightarrow{n \to \infty} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Ilustracja - Mathematica

Prawo wielkich liczb Beroulliego

Twierdzenie

Jeżeli S_n jest liczbą sukcesów w schemacie Bernoulliego n prób z prawdopodobieństwem sukcesu w pojedynczej próbie równym p to dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \leqslant \varepsilon\right) = 1$$

Prawo wielkich liczb Beroulliego

Twierdzenie

Jeżeli S_n jest liczbą sukcesów w schemacie Bernoulliego n prób z prawdopodobieństwem sukcesu w pojedynczej próbie równym p to dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \leqslant \varepsilon\right) = 1$$

Przykład - rzut monetą

Słabe Prawo Wielkich Liczb (Markowa)

Twierdzenie

Niech $(X_n)_n$ będzie ciągiem zmiennych losowych takim, że

- $\lim_{n\to\infty} \frac{\mathcal{D}^2 S_n}{n^2} = 0$
  LUB
- X_n są parami nieskorelowane i mają wspólnie ograniczone wariancje.

wówczas dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n - \mathbb{E}S_n}{n}\right| \geqslant \varepsilon\right) = 0.$$

Słabe Prawo Wielkich Liczb (Markowa)

Twierdzenie

Niech $(X_n)_n$ będzie ciągiem zmiennych losowych takim, że

- $\lim_{n\to\infty} \frac{\mathcal{D}^2 S_n}{n^2} = 0$
  LUB
- X_n są parami nieskorelowane i mają wspólnie ograniczone wariancje.

wówczas dla każdego $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n - \mathbb{E}S_n}{n}\right| \geqslant \varepsilon\right) = 0.$$

Ilustracja - Mathematica

MPWL Bernoulliego

Niech S_n oznacza tym razem liczbe sukcesów w n próbach Bernoulliego z p-stwem sukceu p. Wtedy dla każdego $\varepsilon > 0$ mamy

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{k\geqslant n}\left|\frac{S_n}{k} - p\right| \leqslant \varepsilon\right) = 1.$$

MPWL Bernoulliego

Niech S_n oznacza tym razem liczbe sukcesów w n próbach Bernoulliego z p-stwem sukceu p. Wtedy dla każdego $\varepsilon > 0$ mamy

$$\lim_{n\to\infty} \mathbb{P}\left(\sup_{k\geqslant n}\left|\frac{S_n}{k} - p\right| \leqslant \varepsilon\right) = 1.$$

MPWL Kołmogorowa

Jeżeli $(X_n)_{n\in\mathbb{N}}$ jest ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie $\mathbb{E}|X_1| < \infty$ to $\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}X_1$.

Centralne Twierdzenie Graniczne

Twierdzenie

Niech (X_n) będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie oraz niech $\mathbb{E}X_1 = m$ i $\mathcal{D}^2 X_1 = \sigma^2 > 0$ Wówczas dla dowolnego t

$$\mathbb{P}\left(\frac{X_1 + X_2 + \cdots + X_n - nm}{\sigma\sqrt{n}} \leqslant t\right) \xrightarrow{n\to\infty} \Phi(t).$$

CTG

Twierdzenie

Niech (X_n) będzie ciągiem niezależnych zmiennych losowych o tym samym rozkładzie oraz niech $\mathbb{E}X_1 = m$ i $\mathcal{D}^2 X_1 = \sigma^2 > 0$ Wówczas dla dowolnego t

$$\mathbb{P}\left(\frac{X_1 + X_2 + \cdots + X_n - nm}{\sigma\sqrt{n}} \leqslant t\right) \xrightarrow{n\to\infty} \Phi(t).$$

Ilustracja - Mathematica

Podsumowanie

Dzisiaj rozmawialiśmy o

- testach diagnostycznych raz jeszcze.
- twierdzeniach granicznym.
- Prawach Wielkich Liczb.

Praca domowa

Zadanie 5. [10p]

Podziel plik z wynikami rzutów kostką ($K_1, K_2, \ldots, K_{465}$) na podzbiory

$$X_n = (K_1, K_2, \ldots, K_n), \tag{1}$$

dla $n = 4, 8, 16, 32, 64, 128, 256, 465$. Dla każdego X_n oblicz S_n, a następnie pisząc program dokonujący losowej permutacji danych uzyskaj 100 powtórzeń, dzięki czemu narysujesz histogramy sumy S_n dla $n = 4, 8, 16, 32, 64, 128, 256, 465$ Narysuj osiem histogramów dla każdej z tych wielkości, porównaj z przewidywaniami CTG, a następnie zaprezentuj na wykresie (podobnym do tych z wykładu) zależność S_n od n, skomentuj i zilustruj zgodność z przewidywaniami prawa wielkich liczb (którego?).

W następnym odcinku…

Na następnym wykładzie opowiem o

- Zastosowania MPWL
- Podstawowe metody estymacji punktowej.
- Pożądane właściwości estymatorów.