. . . .
imię, nazwisko, numer indeksu

Egzamin z RPiS (część I, 9 lutego 2023 – 45 minut)

Zadanie 1 (4 punkty)**.** Niech X będzie zmienną losowa taką, ze $P(X = k) = k/10$ dla $k = 1, 2, 3, 4$. Wówczas:

$EX =$ . . . .

$E|X - EX| =$ . . . .

$\mathrm{Var}X =$ . . . .

$g_X(10) =$ . . . .

(g_X to funkcja tworząca prawdopodobieństwa)

Zadanie 2 (3 punkty)**.** Co można powiedzieć o $P(X \geq 10)$ przy następujacych założeniach:

(a) jeśli $X \geq -10, EX = 5$, to $P(X \geq 10) \leq$ . . . .

(b) jeśli $EX = 5$, $\sigma(X) = 2$, to $P(X \geq 10) \leq$ . . . .

(b) jeśli $E(X^4) = 100$, to $P(X \geq 10) \leq$ . . . .

Zadanie 3 (3 punkty)**.** Niech $X_1, \ldots, X_n$ będą zmiennymi losowymi, $x = X_1 + \ldots + X_n$, $\mu = EX$. Które z poniższych założeń są konieczne, by dla dowolnego $\delta > 0$ zachodziła teza nierówności Chernoffa $P(X \geq (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$? Napisz NIE jeśli teza wynika z pozostałych założeń, TAK, jeśli nie wynika (nieprawidłowa odpowiedź w podpunkcie = -1 punkt).

. . . . . . zmienne losowe X_i są niezależne

. . . . . . zmienne losowe X_i mają ten sam rozkład

. . . . . . zmienne losowe X_i są zerojedynkowe

Zadanie 4 (2 punkty)**.** Niech $X_1, X_2, \ldots$ będzie ciągiem *niezależnych* zmiennych losowych o rozkładzie Pois(2), a $Y \sim N(0,1)$. Dla jakiego $f(n)$ zachodzi $\lim_{n\to\infty} P(\sum_{i=1}^{n} X_i > f(n)) = P(Y > 1)$?

$f(n) = \ldots\ldots$

Zadanie 5 (3 punkty)**.** Dla jakiego najmniejszego n istnieje łańcuch Markowa o n stanach i zadanej własności? (Wpisz „nie ma" jeśli nie ma takiego łańcucha Markowa, ∞ lub „nieskończoność" jeśli są takie łańcuchy Markowa, ale mają nieskończenie wiele stanów.)

(a) dwie różne klasy tymczasowe: $n \geq \ldots\ldots$

(b) nie ma stanu powracającego: $n \geq \ldots\ldots$

(c) okresowy: $n \geq \ldots\ldots$

Zadanie 6 (5 punktów)**.** W urnie jest $5-c-z$ kul białych, c kul czerwonych i z kul zielonych. Losujemy (ze zwracaniem) kule, na podstawie wyników losowania próbujemy wyestymować wartości c i z. W kolejnych losowaniach otrzymujemy kulę białą, białą, czerwoną (BBC).

- [1p] funkcja wiarygodności to

  $L(c, z; BBC) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

- [2p] jaką wartość parametrów c i z otrzymamy metodą największej wiarygodności?

  $\hat{c}(BBC) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$
  $\hat{z}(BBC) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

- [1p] Załóżmy, że w n losowaniach wyciągnęliśmy n_b razy kulę białą, n_c razy kulę czerwoną i n_z razy zieloną ($n = n_b + n_c + n_z$). Podaj estymator zgodny i nieobciążony parametru z:

  $\hat{z}(n, n_b, n_c, n_z) = \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

- [1p] Niech z_L, z_R będą takimi statystykami, że $[z_L, z_R]$ jest przedziałem ufności dla parametru z na poziomie 0,9. Czy może się zdarzyć, że (dla pewnych wartości parametrów c i z i pewnej próby) ten przedział nie zawiera żadnej liczby całkowitej? Krótko uzasadnij.

  $\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

  $\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

  $\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$

  $\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$