Przedmiot wyrównawczy
semestr zimowy 2021/2022

Imię i nazwisko: ______________________

zadanie:	2a	2b	Σ:
punkty:	30	30	30
wynik:			
efekty:			

Poniżej przedstawione zostały dwa zadania. Wybierają Państwo jedno z zadań i przedstawiają je do oceny. Na rozwiązanie powinny składać się:
plik `.pdf` - przygotowany w LaTeX lub Markdown plik z opisem rozwiązania, zawierający niezbędną teorię/analizę problemu oraz wnioski płynące z zadania;
plik `.txt` - plik zawierający ewentualny kod programu, jeżeli jest on częścią rozwiązania; kod taki powinien być w dostateczny sposób opisany, aby jasnym było co jest obliczane.

Rozwiązania załączają Państwo w module Zadanie 2. na stronie kursu: plik pdf o nazwie `2_Imie_Nazwisko.pdf` oraz kod programu `2_Imie_Nazwisko.txt` (proszę podać własne imię i nazwisko).

Każde z zadań jest jednakowo punktowane (max. 30 punktów). Zadania rozwiązujemy samodzielnie. Zachęcam do wymiany wskazówek i podejścia do zagadnień, przeglądu literatury czy notatek z rachunku prawdopodobieństwa, itd. na forum kursu, jednak prace, które będą zbyt podobne do siebie w treści otrzymają po 0 punktów.

1. **próbnik Gibbsa, próbkowanie Monte Carlo łańcuchami Markowa** *Zadanie to jest podobne do przykładu generowania próbki z trójwymiarowego rozkładu normalnego. W tym przypadku oczywiście rozkłady brzegowe nie są normalne, a rozkład wektora (X, P, N) nie jest rozkładem ciągłym.*

   Niech zmienne losowe X, P, N mają rozkład łączny:

   $$d(x,p,n) \propto \binom{n}{x} p^x (1-p)^{n-x} \frac{4^n}{n!} \text{ dla } x = 0, 1, \ldots, n,\ 0 < p < 1,\ n = 0, 1, \ldots$$

   (X, N mają rozkłady dyskretne, P jest zmienną losową ciągłą).

   (a) Wyznacz rozkłady warunkowe $(X \mid N = n, P = p)$, $(P \mid X = x, N = n)$, $(N \mid X = x, P = p)$.

   (b) Zapisz algorytm (pseudokod lub kod `R`) generujący liczby pseudolosowe z tego (trójwymiarowego) rozkładu (możesz korzystać z generatorów liczb pseudolosowych z `R` (podstawowych **`runif(), rbeta(), rpois()`**, itd.).

   (c) Przedstaw wykresy rozrzutu (X, P, N), (X, P), (X, N), (N, P) oraz rozkłady brzegowe (histogramy / wykresy masy prawdopodobieństwa X, P i N).

   (d) Zbadaj średnie i wariancje rozkładów brzegowych i porównaj z wartościami teoretycznymi. Zbadaj korelacje/ kowariancje zmiennych X, P, N.

   (e) Porównaj rozkłady warunkowe $(X \mid N, P)$, $(P \mid X, N)$, $(N \mid X, P)$ z symulacji z teoretycznymi (histogram z gęstością teoretyczną, wykres rozkładu masy prawdopodobieństwa dla próbki i teoretyczny).

   Podpunkt (a) wymaga wyznaczenia (analitycznego) teoretycznych rozkładów warunkowych.

   Podpunkt (b) dotyczy algorytmu wykorzystującego próbnik Gibbsa pozwalającego generować liczby pseudolosowe z zadanego rozkładu trójwymiarowego. Znając rozkłady warunkowe $f(\cdot|\diamond)$

z podpunktu (a) możemy generować ciąg zmiennych losowych:

$$(X_0, P_0, N_0), \quad X_1 \sim f(x|P_0 = p_0, N_0 = n_0), \quad P_1 \sim f(p|X_1 = x_1, N_0 = n_0),$$

$$N_1 \sim f(n, |X_1 = x_1, P_1 = p_1), \quad X_2 \sim f(x, |P_1 = p_1, N_1 = n_1), \quad P_2 \sim f(p|X_2 = x_2, N_1 = n_1), \ldots$$

(kolejność losowania kolejnych „współrzędnych” wektora (X, P, N) nie ma znaczenia, wartości początkowe x_0, p_0, n_0 również można ustalić samodzielnie). Generując „na przemian” kolejne wartości, dla dalekich wyrazów generowanego ciągu ($i \to \infty$), wartości te możemy traktować jako próbki z rozkładu d.

2. **łańcuchy Markowa z czasem ciągłym** *Rozważamy stochastyczny model logistycznego wzrostu. Podobnie jak dla modelu SIR wykorzystamy wygenerowane trajektorie, aby określić własności procesu. Rezultaty zestawimy z (deterministycznym) rozwiązaniem odpowiedniego równania różniczkowego zwyczajnego.*

W wersji deterministycznej modelem logistycznego wzrostu nazywamy rozwiązanie $n = n(t)$ równania różniczkowego zwyczajnego

$$\frac{dn}{dt} = rn\left(1 - \frac{n}{K}\right), n(0) = n_0 > 0,$$

gdzie r, K są parametrami.

Interpretacja jest następująca: $n(t)$ opisuje wielkość populacji w chwili t, r oznacza intensywność wzrostu populacji („przyrost naturalny”), a K pojemność środowiska.

W modelu stochastycznym modelujemy wprost „każdego osobnika” w populacji; wielkość populacji $n(t)$ określa liczbę osobników, z których każdy może się rozmnożyć ($n \to n+1$) lub umrzeć ($n \to n-1$). Załóżmy, że w chwili t mamy $n(t) = i \geqslant 0$ osobników oraz, że z intensywnością λ_i w populacji pojawia się nowy osobnik ($i \to i+1$), a z intensywnością μ_i jeden z osobników umiera ($i \to i-1$). Oczywiście $\lambda_0 = \mu_0 = 0$.

Stochastyczny model logistyczny powinien spełniać następujące warunki:

$$\lambda_n - \mu_n = rn - \frac{r}{K}n^2,$$
$$\lambda_K = \mu_K.$$

Rozważmy dwa różne przypadki spełniające powyższe założenia:

(i) $\lambda_i = \begin{cases} i - \frac{i^2}{100} & i = 0, 1, \ldots, 100, \\ 0 & i > 100, \end{cases}, \quad \mu_i = \frac{i^2}{100}, i = 1, 2, \ldots$

(ii) $\lambda_i = i, \mu_i = \frac{i^2}{50}, i = 0, 1, 2, \ldots$.

Przyjmijmy, że $n_0 = 5$.

(a) Dla modelu deterministycznego, przedstaw rozwiązanie numeryczne $n(t)$ (pakiet `deSolve`), dla różnych wartości parametrów r, K oraz różnych warunków początkowych n_0.

(b) Następnie dla modelu stochastycznego: wyznacz wartość oczekiwaną $E[n(t)]$ oraz wariancję $var[n(t)]$ dla ustalonej chwili np. $t = 10$;

(c) wyznacz funkcję wartości oczekiwanej $E[n(t)]$ oraz wariancji $var[n(t)]$, $t \geqslant 0$.

(d) Porównaj $E[n(t)]$ z rozwiązaniem deterministycznym.

(e) Opisz różnice pomiędzy modelem deterministycznym oraz modelem stochastycznym (oraz między wersjami (i) oraz (ii)).