

Examen PMP (29.1.2024)
Test 2
 timp de lucru: 50 min.

Răspunsul de la teste va trebui încărcat, până la finalul testului, pe “repozitorul” personal de pe GitHub, într-un director numit “Examen”. Pe lângă codul în Python, vor trebui încărcate și screenshoturi ale execuțiilor, precum și eventualele grafice.

Subiectul 1. Preziceți dacă un pasager de pe Titanic a supraviețuit sau nu pe baza clasei de pasageri și a vârstei.

- Încărcați setul de date Titanic.csv într-un Pandas DataFrame și preprocesați datele prin gestionarea valorilor lipsă, transformarea variabilelor (dacă este necesar), etc.
- Definiți modelul în PyMC folosind cele două variabile independente (clasa de pasageri: *Pclass* și vârsta: *Age*) pentru a prezice variabila dependentă (*Survived*).
- Care credeți că este variabila care influențează cel mai mult rezultatul (dacă pasagerul a supraviețuit sau nu)?
- Construiți un interval de 90% HDI pentru probabilitatea de a supraviețui a unui pasager de 30 ani de la clasa a 2^a.

Subiectul 2. Fie X și Y variabile aleatoare independente repartizate Geometric cu parametrii 0.3 și 0.5, respectiv.

- Aplicați metoda Monte Carlo cu 10000 de iterații pentru a aproxima $P(X > Y^2)$.
Indicație: Puteți să vă inspirați din codul de la pag. 21 din cursul 11 (ce aproximează numărul π), condiția de a fi în interiorul cercului fiind înlocuită de cea ca $x > y^2$. O simulare a unei variabile aleatoare repartizate Geometric(θ) se poate obține, de exemplu cu `stats.geom.rvs(θ)` în SciPy.
- Neavând la dispoziție valoarea căutată, veți determina o medie pentru $k = 30$ astfel de aproximații și deviația standard corespunzătoare.

Tabel punctaj:

subiect:	1.a.	1.b.	1.c.	1.d.	2.a.	2.b.	oficiu	Total
punctaj max.:	2	7	4	6	7	3	1	30
punctaj obținut:								