LTTC-mock

Choose four out of the five questions.

Probability spaces and the generation of random variables

Define an explicit function $\phi : [0,1] \to [0,1] \times \{1,2\}$ such that:

- If U is uniformly distributed in $[0,1]$, then $\phi(U)$ has the same distribution as (V, X), where V and X are independent and uniformly distributed on $[0,1]$ and $\{1,2\}$, respectively;
- There exists a set $A \subset [0,1]$, with $\mathbb{P}(U \in A) = 1$, such that ϕ is injective on A.

Solution

Let $b : [0,1] \to \{0,1\}^{\mathbb{Z}^+}$ be so that $b(u)$ is the binary expansion of $u \in [0,1]$ this is bijective modulo a countable set. We know that $b(U)$ is a sequence of iid Benroulli $\frac{1}{2}$ random variables, from which we define

$$\phi(u) = (b^{-1}[b(u)_{n=2}^{\infty}], b(u)_1).$$

Borel-Cantelli

Let $(X_i)_{i=2}^{\infty}$ be i.i.d. continuous random variables with probability density function given by

$$f(x) = \frac{1}{2}|x|^3 e^{-x^4/4}.$$

Find a deterministic sequence $(c_i)_{i=2}^{\infty}$ such that almost surely

$$\limsup_{n\to\infty} \frac{X_n}{c_n} = 1.$$

Solution

For $x \geq 0$, we easily see that

$$\mathbb{P}(X_n > x) = \frac{1}{2}e^{-x^4/4}.$$

Take

$$c_n = (4\log n)^{\frac{1}{4}}.$$

Will we show using the first Borel-Cantelli lemma that almost surely,

$$\limsup_{n\to\infty} \frac{X_n}{c_n} \leq 1$$

and then we will show using the second Borel-Cantelli lemma and the independence of the X_i that almost surely

$$\limsup_{n\to\infty} \frac{X_n}{c_n} \geq 1.$$

Let $\epsilon \geq 0$. We have that

$$\mathbb{P}(X_n/c_n > 1+\epsilon) = \frac{1}{2n^{(1+\epsilon)^4}} =: a_n.$$

The sequence a_n is summable for $\epsilon > 0$, so the upper bound follows by from the first Borel-Cantelli lemma. For the lower bound, if we set $\epsilon = 0$, then a_n is not summable, and then the lower bound follows from the second Borel-Cantelli lemma.

Markov chains

Prove that for an irreducible aperiodic Markov chain on a finite state space S, we have that for each $s \in S$, the return time

$$T := \inf\{n \geq 1 : X_n = s\}$$

has finite expectation, regardless of the starting distribution of the chain.

Solution

We go back to basics; in particular, some observations that were made in connection to the Doeblin coupling. If P is the transition matrix, then by the assumptions of irreducibility and aperiodicity, we know that P^M has all positive entries for some $M > 0$; let $\delta > 0$ be the smallest entry. Hence every M steps, we have a non-zero probability $\delta > 0$ of getting to the state s. Thus

$$\mathbb{P}(T > kM) \leq (1-\delta)^k$$

from which we easily deduce that T has finite expectation.

Poisson processes

Suppose that Π is a Poisson process on $\mathbb{R}^d$. Let $t \in \mathbb{R}^d$, and $c \in (0, \infty)$ Prove that:

- The translated point process $\Gamma := t + \Pi$ given by translating each point of Π by t is still a Poisson point process on $\mathbb{R}^d$.
- The scaled point process $\Sigma := c\Pi$ given by multiplying each point of Π by c is still a Poisson point process on $\mathbb{R}^d$.

Solution

Suppose that Π is a Poisson point process of intensity λ. Then Π satisfies:

- $\Pi(A)$ is a Poisson random variable with mean $\lambda|A|$ for every set A of finite volume.
- The random variables $\Pi(A_1), \ldots, \Pi(A_n)$ are independent for pairwise disjoint sets $A_1, \ldots, A_n$.

It suffices to verify these properties for Γ and Σ. For a subset $A \subset \mathbb{R}^d$, let $t + A = \{t + a : a \in A\}$ and $cA = \{ca : a \in A\}$.

- Observe that $\Gamma(A) = \Pi(-t + A)$, so that the two required properties are easily verified.
- Similarly $\Sigma(A) = \Pi(c^{-1}A)$, and Σ is a Poisson point process of intensity $c^{-1}\lambda$.

Estimating the stationary distribution

- Suppose you are given the output of a 100000 steps of a irreducible and aperiodic finite state Markov chain. Carefully explain how you could estimate the stationary distribution for this Markov chain, and why you estimator is reasonable.
- Import the data from the file markovchain.txt (https://tsoo-math.github.io/ucl2/markovchain.txt) and use this data and your method above to estimate the stationary distribution.

Solution

- Since the chain is on a finite state space and irreducible and aperiodic, it has a unique stationary distribution π, and we know that a version of the law of large numbers gives that for every state s, we have

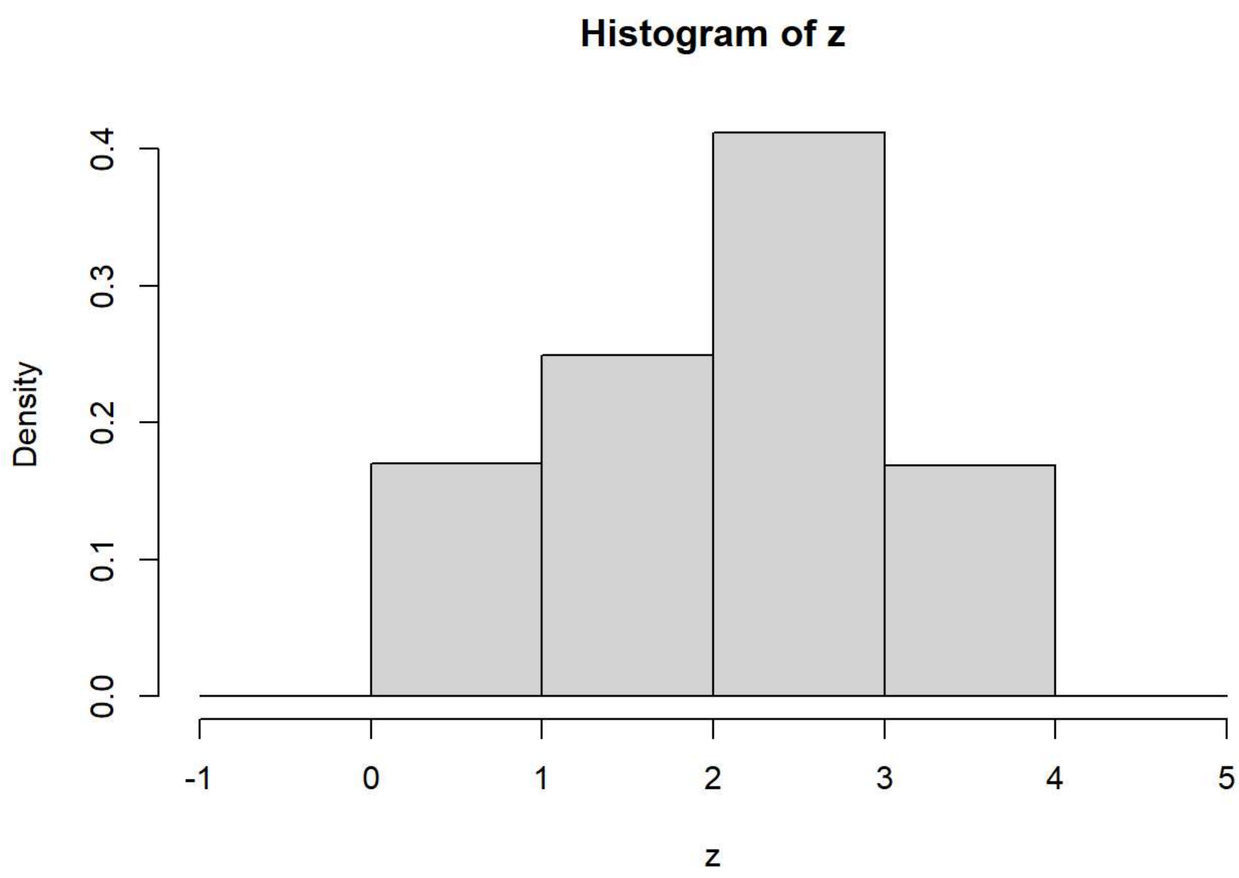
$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}[X_i = s] \to \pi(s).$$

Thus for each state, we simply need to count the number of occurrences and divide by 100000, to get its approximate probability under π.

- A quick scan of the file shows there are 4 states: $1, 2, 3, 4$. With R, we have:

```
z =read.table("markovchain.txt", sep=",")
z= z[,]

b1 = seq(-1,5, by=1)
hist(z, prob=TRUE, breaks=b1)
```



```
sum(z==1)/100000
```

```
## [1] 0.17023
```

```
sum(z==2)/100000
```

```
## [1] 0.24957
```

```
sum(z==3)/100000
```

```
## [1] 0.41172
```

```
sum(z==4)/100000
```

```
## [1] 0.16849
```