Technical University of Munich

Detection of trail of evidence in a graph

Adam Menoud

Thesis submitted
within the Bachelors's Program
Mathematics

Supervisors: Prof. Noam Berger, PhD Guilherme de Paula Reis

September 21, 2023

I hereby declare that this thesis is my own work and that no other sources have been used except those clearly indicated and referenced.

Thesis advisor: Prof. Noam Berger, PhD Guilherme de Paula Reis *Adam Menoud*

Detection of trail of evidence in a graph

ABSTRACT

In this paper, we consider Gaussian random variables indexed by the nodes of a graph, and try to asymptotically choose between two hypotheses; namely whether there exists a path which is ”elevated” in the sense that the corresponding random variables have expectation higher than 0. We consider two different graphs, the regular lattice of dimension 2 and the complete binary tree. In both cases, we consider both the minimax and the bayesian approach, and try to determine if determining the correct hypothesis is asymptotically possible for different decaying rate of the elevation; this will take the form of 3 theorems discussing every different cases. All the following is a development of (a part of) the article ”Searching for a trail of evidence in a maze” by Arias-Castro et al. , with the intention of making it clearer and more understandable for less experienced mathematics students.

Contents

Chapter 1

Introduction:

1.1 Problem setup and first definitions.

For this project, we will deal with the following problem.
We consider a graph $G = (V, E)$, more specifically either a regular lattice or a complete binary tree. We also have a set of independent normal random variables $\{X_\nu, \nu \in V\}$ that correspond to each vertex of the graph.
We want to be able to distinguish between two cases: either all those random variables are equally distributed with mean 0, or there exists a path such that the mean is equally higher for all the random variables on the path.
Note that we require that such a path has a specific shape, in a certain way "natural" with respect to the graph considered.
Let us define precisely the graphs and the corresponding "valid" paths.

Definition:
The (2-dimensional) **regular lattice** of size m is the graph whose vertices are given by $V_m := \{(i,j) \in \mathbb{Z}^2; 0 \leq i \leq m-1, |j| \leq i, j \equiv i \mod 2\}$ with the (oriented) edges $E_m := \{((i,j);(i,j\pm 1)); 0 \leq i \leq m-1\}$

m refers here to the size of the problem; as it will be explained later, our objective will be to decide if it is possible to detect the presence of a path as $m \to \infty$.
Throughout the whole document, when the variable is not precised, we **mean that limits are taken as** $m \to \infty$.

Notice that the regular lattice above could be defined on the set $\mathbb{N}^2$; however, the shape that we chose will be more convenient for later (since it resembles the random walks with i as index).
The valid paths for this graph are the ones starting at the root $(0,0)$ and following the directed edges until a node $(m-1,k)$, at the far right of the graph. We denote the set of all those paths as P_m.
If we generate such a path by successively picking its vertices, since we do not choose the root and then we have $(m-1)$ binary choices to make, we have $|P_m| = 2^{m-1}$.

For the second type of graph, we will denote its component with the exact same notations; as the rest of the document is clearly separated into parts involving either the regular lattice or the binary tree, it should never be a problem.

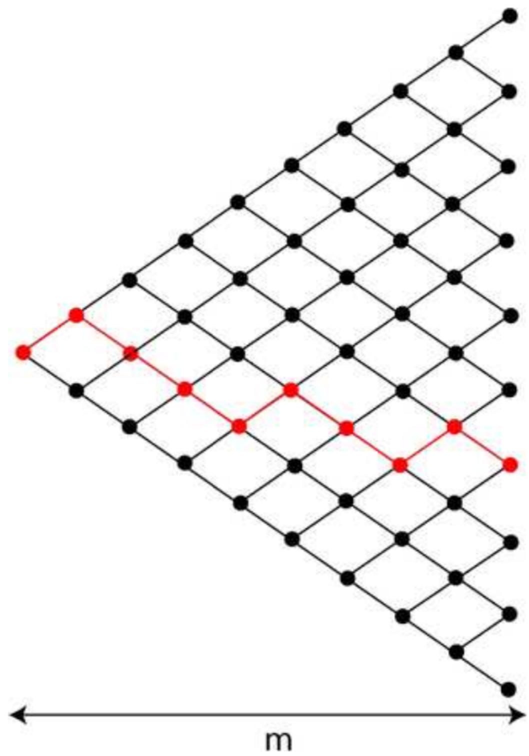

Figure 1.1: A path in the regular lattice

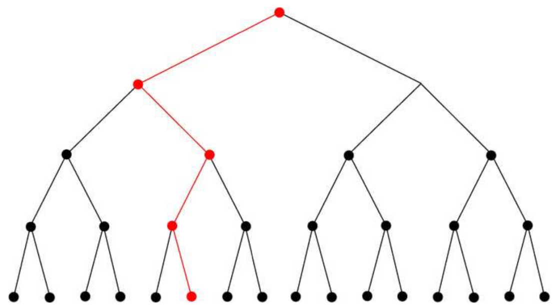

Figure 1.2: A path in (an isomorphism of) the binary tree

Definition: the (complete) **binary tree** of size m if the graph whose vertices are given by $V_m := \{(i,j) \in \mathbb{Z}^2; 0 \leq i \leq m-1, 0 \leq j < 2^i\}$ and (oriented) edges by $E_m = \{((i,j),(i+1,2j+b)); 0 \leq i \leq m-1, b \in \{0,1\}\}$

The set of paths P_m is defined exactly as before, from the root $(0,0)$, and the exact same reasoning also leads to $|P_m| = 2^{m-1}$. We also notice (by summing all the columns) that $|V_m| = 1 + 2 + 4 + ... + 2^{m-1} = 2^m - 1$

With the precise definitions for the graphs in mind, we can further define the statement of the problem.
The following is valid both is the case of the regular lattice and the case of the binary tree.
We name H_0 the simple hypothesis where all the random variables corresponding to the vertices are standard normal distributed, i.e.

$$\forall \nu \in V_m, X_\nu \sim \mathcal{N}(0,1)$$

For a path $p \in P_m$, we name $H_{1,p}$ the simple hypothesis where all the random variables along the path are elevated by some number $\mu_m > 0$, i.e

$$\forall \nu \in V_m, \quad X_\nu \sim \begin{cases} \mathcal{N}(0,1) & \text{if } \nu \notin p \\ \mathcal{N}(\mu_m,1) & \text{if } \nu \in p \end{cases}$$

Finally, H_1 is the name of the hypothesis being true if and only if there exists a path $p \in P_m$ such that $H_{1,p}$ is true. It is no longer a simple hypothesis, since it contains 2^{m-1} distributions.

For the different hypothesis above, we denote the **probability of an event** A **under the hypothesis** $H_0, H_{1,p}$ or H_1 respectively by $\mathbb{P}_0(A), \mathbb{P}_{1,p}(A)$ and $\mathbb{P}_1(A)$. We also define

$$\mathbb{P}_\pi(A) := \mathbb{E}_\pi(\mathbb{P}_{1,p}(A)) = 2^{1-m} \sum_{p \in P_m} \mathbb{P}_{1,p}(A)$$

the average distribution given by the uniform prior on paths π.

Furthermore, to try to discriminate between hypotheses H_0 and H_1, we use a test:
Definition: A **test** is a a function T from the data with binary outcome:

$$T : \{X_\nu, \nu \in V_m\} \to \{0, 1\}$$

Such a test should output 0 when H_0 is true, and 1 when H_1 is true.

For all the following, such a test is always of the form $\mathbb{I}\{S \in C\}$, where S is a statistic and $C \subseteq \mathbb{R}$ is called the critical region (C is always a half-line in our case, even though it could be more general). That explains why we will most often focus on a statistic together with a threshold, considered as a test.

As we have seen before, in our case, H_1 is composed of many different cases; there is two useful ways to reduce the composite hypothesis H_1 into a simple one.

The first is the bayesian approach; we choose a prior distribution π on the discrete probability space $(P_m, \mathcal{P}(P_m))$ and we suppose that if H_1 were to be true, $H_{1,p}$ would be true with probability $\pi(p)$.
Since we do not have any information about this prior distribution, we will assume that it is the uniform one, i.e.

$$\forall p \in P_m, \pi(p) = 2^{-(m-1)}$$

With this idea in mind, we can define the bayesian risk of a test T (given a prior distribution π):

$$\gamma_\pi(T) := \mathbb{P}_0(T = 1) + 2^{-(m-1)} \sum_{p \in P_m} \mathbb{P}_{1,p}(T = 0)$$

The other way to consider the problem is the minimax approach. Here, instead of taking the average over the different outcomes, we consider the case $(H_{1,p})$ versus H_0 one by one, and we only consider the higher risk. This leads to the following definition of the maximum (or minimax) risk:

$$\gamma(T) := \mathbb{P}_0(T = 1) + \sup_{p \in P_m} \mathbb{P}_{1,p}(T = 0)$$

Finally, we define (for a test T) the **probability of Type I error** as $\mathbb{P}_0(T = 1)$, and the **probability of Type II error** as $2^{-(m-1)} \sum_{p \in P_m} \mathbb{P}_{1,p}(T = 0)$ for the bayesian case and as $\sup_{p \in P_m} \mathbb{P}_{1,p}(T = 0)$ for the minimax case.

1.2 Statement of the theorems

All theorems give us a condition on the elevation μ_m for the path to be either detectable or undetectable. For the two first results, we consider the case of the (2-dimensional) regular lattice:

1.2.1 Theorem 1.1

We consider the minimax risk γ.

(a) Suppose $\mu_m\sqrt{\log m} \to \infty$
Then there is a sequence (T_m) of tests such that $\gamma(T_m) \to 0$
We say that the sequence of test is **(asymptotically) powerfull** (with respect to the minimax risk).
(b) Suppose $\mu_m \log m\sqrt{\log\log m} \to 0$
Then for any sequence of test (T_m),we have $\limsup_{m\to\infty} \gamma(T_m) \geq 1$
We say that the sequence of test is **(asymptotically) powerless** (with respect to the minimax risk).

We can notice that in some cases, this theorem cannot conclude anything.

1.2.2 Theorem 1.2

We obtain similar results for the bayesian risk γ_π (with π the uniform distribution on P_m):
(a) Suppose $\mu_m \log(m)^{-1/4} \to \infty$
Then there is a sequence (T_m) of tests such that $\gamma_\pi(T_m) \to 0$
(i.e. it is powerful with respect to the bayesian risk)
(b) Suppose $\mu_m \log(m)^{-1/4} \to 0$
Then any sequence of test (T_m) is asymptotically powerless (with respect to the bayesian risk)

We will now consider the case of the complete binary tree. Since we obtain the same results for the minimax and the bayesian risk, there is only one theorem about this graph.

1.2.3 Theorem 1.3

(a) Suppose $\mu = \mu_m \geq \sqrt{2\log 2}$
Then there exist a sequence of test $(T_m)_{m\geq 1}$ so that $\gamma(T_m) \to 0$
Note that this also implies $\gamma_\pi(T_m) \to 0$
(b) Suppose $\mu = \mu_m < \sqrt{2\log 2}$
Then there exist no powerful sequence of test (with respect to both risks).
(c) If $\mu_m \to 0$, then any sequence of test is powerless (with respect to both risks).

It is important to notice that ”powerless” is not the same as ”not powerful”, but rather a special case. A sequence of test $(T_m)_{m\geq 1}$ is not powerful if

$$\liminf_{m\to\infty} \gamma(T_m) =: l > 0$$

and it is powerless if $l = 1$.

1.3 The likelyhood ratio

Before proving the theorems, we will compute the likelihood ration and its expectation and variance under H_0, since those results are used several times among the three theorems.
The calculations are true for paths both in the regular lattice and the binary tree.

For $p \in P_m$, we define $X_p := \sum_{\nu\in p} X_\nu$
($\nu \in p$ means that the path p goes through the node ν) and $f_{Y,H}$ is the density of a random element Y under the hypothesis H. X is the random element given by every node of the graph.
We first compute the likelihood ratio, defined as follows:

$$\begin{aligned}
L_m(X_\nu, \nu \in V_m) &:= \frac{d\mathbb{P}_\pi(X)}{d\mathbb{P}_0(X)} \\
&= \frac{2^{-(m-1)} \sum_{p\in P_m} f_{X,H_{1,p}}(X_\nu, \nu \in V_m)}{f_{X,H_0}(X_\nu, \nu \in V_m)} \\
&= \frac{2^{-(m-1)} \sum_{p\in P_m} \prod_{\nu\in V_m} f_{X_\nu,H_{1,p}}(X_\nu)}{\prod_{\nu\in V_m} f_{X_\nu,H_0}(X_\nu)} \\
&= 2^{-(m-1)} \sum_{p\in P_m} \frac{\prod_{\nu\in p} f_{X_\nu,H_{1,p}}(X_\nu)}{\prod_{\nu\in V} f_{X_\nu,H_0}(X_\nu)} \\
&= 2^{-(m-1)} \sum_{p\in P_m} \prod_{\nu\in p} \frac{\exp(-\frac{1}{2}(X_\nu - \mu_m)^2)}{\exp(-\frac{1}{2}(X_\nu)^2)} \\
&= 2^{-(m-1)} \sum_{p\in P_m} \prod_{\nu\in p} \exp(X_\nu \mu_m - \frac{\mu_m^2}{2}) \\
&= 2^{-(m-1)} \sum_{p\in P_m} \exp(\sum_{\nu\in p}(X_\nu \mu_m - \frac{\mu_m^2}{2})) \\
&= 2^{-(m-1)} \sum_{p\in P_m} \exp(X_p \mu_m - \frac{m\mu_m^2}{2})
\end{aligned}$$

Now, we are interested in computing the expectation and the variance of this ratio under the hypothesis H_0 (we denote it by $\mathbb{E}_0$ and Var_0).

Expectation:

$$
\begin{aligned}
\mathbb{E}_0 L_m(X_\nu, \nu \in V_m) &= \mathbb{E}_0 \; 2^{-(m-1)} \sum_{p \in P_m} \exp(X_p \mu_m - \frac{m\mu_m^2}{2}) \\
&= 2^{-(m-1)} \sum_{p \in P_m} \mathbb{E}_0 \exp(X_p \mu_m - \frac{m\mu_m^2}{2}) \\
&= 2^{-(m-1)} e^{-\frac{m\mu_m^2}{2}} \sum_{p \in P_m} \mathbb{E}_0 \exp(X_p \mu_m) \\
&= 2^{-(m-1)} e^{-\frac{m\mu_m^2}{2}} \sum_{p \in P_m} e^{\frac{m\mu_m^2}{2}} \\
&= 1
\end{aligned}
$$

We used the formula for the MGF of a normal random variable for the 4th step.

Variance: We abuse the notation and note $\nu \in p$ if the path p goes through the vertex ν, and $\nu \in p \cap q$ if this is the case for both path p and q.
For the following, we define:
For two paths $p, q \in P_m$, $N_{p,q} := |\{\nu \in V_m, \nu \in p \cap q\}|$
And $N_m : (P_m^2, \mathcal{P}(P_m^2), \pi \otimes \pi) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ as the number of crossings between two paths; it is a discrete random variable with value in $\{1, ..., m\}$

With all of these definitions, we compute the variance (under the hypothesis H_0 ; we write Var_0 for the variance, and Cov_0 for the covariance)

$$
\begin{aligned}
Var_0(L_m) &= Var_0(\frac{1}{2^{m-1}}\sum_{p\in P_m}\exp(\mu_m X_p - \frac{m\mu_m^2}{2})) \\
&= (\frac{e^{-\frac{m\mu_m^2}{2}}}{2^{m-1}})^2 Var_0(\sum_{p\in P_m} e^{\mu_m X_p}) \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} Var_0(\sum_{p\in P_m} e^{\mu_m X_p}) \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} Cov_0(e^{\mu_m X_p}, e^{\mu_m X_q}) \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} \mathbb{E}_0\left[(e^{\mu_m X_p} - \mathbb{E}_0(e^{\mu_m X_p})(e^{\mu_m X_q} - \mathbb{E}_0(e^{\mu_m X_q})\right] \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} \mathbb{E}_0\left[e^{\mu_m(X_p+X_q)} - \mathbb{E}_0(e^{\mu_m X_p})e^{\mu_m X_q} - \mathbb{E}_0(e^{\mu_m X_q})e^{\mu_m X_p} + \mathbb{E}_0(e^{\mu_m X_q})\mathbb{E}_0(e^{\mu_m X_p})\right] \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} \mathbb{E}_0\left[e^{\mu_m(X_p+X_q)}\right] - \mathbb{E}_0(e^{\mu_m X_p}) \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} \mathbb{E}_0(e^{2\mu_m\sum_{\nu\in p\cap q}X_\nu})\mathbb{E}_0(e^{\mu_m\sum_{\nu\notin p\cap q}X_\nu}) - e^{\frac{m\mu_m^2}{2}} \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} e^{2\mu_m^2 N_{p,q}} e^{\mu_m^2(m-N_{p,q})} - e^{\frac{m\mu_m^2}{2}} \\
&= \frac{e^{-m\mu_m^2}}{2^{2(m-1)}} \sum_{p,q\in P_m} e^{N_{p,q}+m} - e^{\frac{m\mu_m^2}{2}} \\
&= \frac{1}{2^{2(m-1)}} \sum_{p,q\in P_m} e^{\mu_m^2 N_{p,q}} - 1 \\
&= \mathbb{E}\left[e^{\mu_m^2 N_m}\right] - 1
\end{aligned}
$$

optimality: Lemma 1.1: The likelihood ratio test with threshold 1 is minimal with respect to the baysian risk.
Proof: Denote by f_0 and $f_\pi = 2^{1-m}\sum_{p\in P_m} f_{X,H_{1,p}}$ the density of $\mathbb{P}_0$ and $\mathbb{P}_\pi$ on $\mathbb{R}^{|V_m|}$ (with respect to the Lebesgue measure). For a test T, we have

$$
\begin{aligned}
\gamma_\pi(T) &= \mathbb{P}_0(T=1) + \mathbb{P}_\pi(T=0) \\
&= \int_{\mathbb{R}^{|V_m|}} f_0(x)\mathbb{I}\{T(x)=1\} + f_\pi(x)\mathbb{I}\{T(x)=0\}dx
\end{aligned}
$$

To minimize the integrand (and thus the integral), we choose T so that $T=0$ when $f_0(x) > f_\pi(x)$, $T=1$ when $f_0(x) < f_\pi(x)$, and either 0 or 1 in case of equality. Hence, a minimizing test in given by

$$T(X) := \mathbb{I}\{f_0(X) < f_\pi(X)\} = \mathbb{I}\{\frac{f_\pi(X)}{f_0(X)} > 1\}$$

□

We will use this result of optimality for several proofs of this paper (in fact, every time that we want to prove that any sequence of test is powerless).
Also notice that since the minimax risk is by definition greather or equal than the baysian risk, we will also be able to use this result in the minimax case.

With all the previous calculations, we are equipped to prove the theorems.

Chapter 2

Proof theorem 1.1: Regular lattice, minimax risk

2.1 Part (a): Powerful test

The statistic we will use for the test is the weighted sum

$$S_m := \sum_{(i,j)\in V_m} \omega_i X_{i,j}, \quad \omega_i := \frac{\lambda_m}{i+1} \quad \lambda_m := (\sum_{i=0}^{m-1} \frac{1}{i+1})^{-1}$$

The constant λ_m is constructed so that the sum of all weights ω_i is 1.
By definition of S_m, we have that its distribution under H_0 is $\mathcal{N}(0, \lambda_m)$ and it is $\mathcal{N}(\mu_m, \lambda_m)$ under H_1.

We use the test that reject the null if $S_m > \frac{\mu_m}{2}$
Now it simply suffices to compute the (minimax) risk:
(Z is here an independent standard Gaussian random variable)

$$\begin{aligned}
\gamma(\mathbb{I}\{S_m > \frac{\mu_m}{2}\}) &= \mathbb{P}_0(S_m > \frac{\mu_m}{2}) + \sup_{p\in P_m} \mathbb{P}_{1,p}(S_m \leq \frac{\mu_m}{2}) \\
&= \mathbb{P}(\sqrt{(\lambda_m)}Z > \frac{\mu_m}{2}) + \mathbb{P}(\sqrt{(\lambda_m)}Z + \lambda_m \leq \frac{\mu_m}{2}) \\
&= \mathbb{P}(Z > \frac{\mu_m}{2\sqrt{(\lambda_m)}}) + \mathbb{P}(Z \leq \frac{-\mu_m}{2\sqrt{(\lambda_m)}}) \\
&= \mathbb{P}(Z \geq \frac{\mu_m}{2\sqrt{\lambda_m}})
\end{aligned}$$

and we see that this quantity goes to 0 if $\frac{\mu_m}{\sqrt{\lambda_m}} \to \infty$
This condition is indeed fulfilled, since it follows from definition that λ_m has the asymptotic behavior of $(\log m)^{-1}$, and by assumption we have $\mu_m\sqrt{\log m} \to \infty$
$\square$
For the following proof, we will use some results about what is called the "predictability profile", defined as follows:
Definition: A **nearest-neightbor** process is a discrete random process $S = (S_m)_{m\in\mathbb{N}}$ (with $S_0 = 0$) so that for all $m \in \mathbb{N}, S_{m+1} \in \{S_m + 1, S_m - 1\}$)

Definition: For $S = (S_m)_{m\geq 1}$ a nearest-neightbor process , the **predictability profile** of S is defined to be

$$PRE_S(k) := \sup_{n\in\mathbb{N}, x\in\mathbb{Z}, (S_i)_{i=0}^n \ NN-process} \mathbb{P}(S_{n+k} = x | S_0, ..., S_n)$$

(The (truncated) nearest-neighbor process $(S_i)_{i=0}^n$ should here be considered as "a history"). In plain English, the predictability profile give for any k the maximal (among all the step of the random walk) probability of guessing k steps into the future, with knowledge about the previous steps.

The two lemmas that we will use are the following:
Lemma 2.1 (lemma 2.3 in [2]): Take $(a_j)j \in \mathbb{N}$ such that $\sum_{j\in\ \mathbb{N}} a_j < 1$. Then there exists a nearest-neighbor process $S = (S_k)_{k\geq 0}$ so that $\forall k \geq 1$

$$PRE_S(k) \leq \frac{20}{k a_{\lfloor \log_2 \frac{k}{2} \rfloor}}$$

Lemma 2.2 (lemma 3.1 in [3]): Consider $S = (S_k)_{k=1}^{m-1}$
Take a real number B so that $\sum_{k=1}^{\lfloor \frac{m}{B} \rfloor} PRE_S(kB) \leq \theta < 1$
For any sequence $v = (v_n)_{0\leq n\leq m-1}$ so that $(n, v_n) \in V_m$ (so in particular for a path in P_m in our case), we have

$$\mathbb{P}(|S \cap v| \geq k) \leq \theta^{k/B} B$$

where $S \cap v := \{n,\ S_n = v_n\}$ is the set of indices for which S overlaps with v.

2.2 Part (b): All tests are powerless

Proof: From section 1.4, we have

$$B_m(\pi) := \inf_{tests\ T} \gamma_\pi(T) = \mathbb{P}_0(L_m \geq 1) + \mathbb{P}_\pi(L_m < 1)$$

We also claim:

$$B_m(\pi) = 1 - \frac{\mathbb{E}_0|L_m - 1|}{2} \geq 1 - \frac{\sqrt{\mathbb{E}_0(L_m-1)^2}}{2} \qquad (2.1)$$

The inequality follows from Cauchy's inequality, and we show the first equality. First, notice that for an event A, we have

$$\mathbb{E}_0(L_m \mathbb{I}_A) = \int \mathbb{I}_A \frac{d\mathbb{P}_\pi}{d\mathbb{P}_0} d\mathbb{P}_0 = \mathbb{P}_\pi(A)$$

With that, we compute:

$$\begin{aligned}
\mathbb{E}_0(|L_m - 1|) &= \mathbb{E}_0((L_m-1)\mathbb{I}\{L_m \geq 1\}) - \mathbb{E}_0((L_m-1)\mathbb{I}\{L_m < 1\}) \\
&= \mathbb{P}_\pi(L_m \geq 1) - \mathbb{P}_0(L_m \geq 1) - \mathbb{P}_\pi(L_m < 1) + \mathbb{P}_0(L_m < 1) \\
&= 2 - 2\mathbb{P}_0(L_m \geq 1) - 2\mathbb{P}_\pi(L_m < 1)
\end{aligned}$$

And we prove the claim by rearranging the terms.

The inequality (2.1) implies that if we can show $\mathbb{E}_0(L_m - 1)^2 \to 0$, then any test is asymptotically powerless and so we are done.
The calculation of the variance of L_m (see section 1.4) shows

$$\mathbb{E}_0(L_m - 1)^2 = \mathbb{E}e^{\mu_m^2 N_m} - 1$$

with $N_m : (P_m^2, 2^{P_m^2}, \pi^2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ the number of crossings between two paths.
By combining it with inequality (2.1), we see that if $\mathbb{E}e^{\mu_m^2 N_m} \to 1$, then the risk is bounded below by 1, implying that any test is asymptotically powerless.
In order to use the lemmas above, we define (for $m \geq 2$)

$$a_j = a_j(m) := \frac{1}{3\log_2 m}\mathbb{I}\{j \leq \log_2 m\}$$

We check that the sum is less than 1 (for $m > 4$):

$$\sum_{j \in \mathbb{N}} a_j = \frac{\lfloor \log_2 m \rfloor + 1}{3\log_2 m} < \frac{1}{2} < 1$$

With that choice, the lemma 2.1 gives us the following inequality, for any $k \geq 1$ and some nearest-neighbor process S:

$$PRE_S(k) \leq \frac{20}{ka_{\lfloor \log_2 \frac{k}{2} \rfloor}} = \frac{20}{k\mathbb{I}\{k \leq m\}(\frac{1}{3\log_2 m})} = \frac{60\log_2 m}{k}$$

since we have $k \leq m$. By defining $B = B_m := \frac{120(\log m)^2}{\log 2}$, we have

$$\begin{aligned}
\sum_{k=1}^{\lfloor \frac{m}{B_m} \rfloor} PRE_S(kB_m) &= \sum_{k=1}^{\lfloor \frac{m}{B_m} \rfloor} \frac{60\log_2 m \log 2}{k120(\log m)^2} \\
&= \frac{1}{2\log m} \sum_{k=1}^{\lfloor \frac{m}{B_m} \rfloor} \frac{1}{k} \\
&\leq \frac{1}{2\log m} \log(\lfloor \frac{m}{B_m} \rfloor - 1) \\
&\leq \frac{1}{2}
\end{aligned}$$

Using the previous result, we can apply lemma 2.2 to deduce that

$$\mathbb{P}(N_m \geq k) \leq 2^{\frac{-k}{B_m}} B_m$$

since N_m is defined to be the number of times a path (equivalently, a nearest-neighbor process) $p \in P_m$ intersects with another path $q \in P_m$.

With that result in mind, we now compute the interesting quantity; for the second step, we use the following summation by part formula:

$$\sum_{k=K}^{n} f_k(g_k - g_{k+1}) = \sum_{k=K+1}^{n} g_k(f_k - f_{k-1}) + f_K g_K - f_n g_{n+1}$$

Which gives us the following:

$$\begin{aligned}
\mathbb{E}e^{\mu_m^2 N_m} &= \sum_{k=1}^{K-1} e^{\mu_m^2 k}\mathbb{P}(N_m = k) + \sum_{k=K}^{\infty} e^{\mu_m^2 k}\left[\mathbb{P}(N_m \geq k) - \mathbb{P}(N_m \geq k+1)\right] \\
&= \sum_{k=1}^{K-1} e^{\mu_m^2 k}\mathbb{P}(N_m = k) + \sum_{k=K+1}^{\infty} \mathbb{P}(N_m \geq k)\left[e^{\mu_m^2 k} - e^{\mu_m^2 (k-1)}\right] \\
&+ e^{\mu_m^2 K}\mathbb{P}(N_m \geq K) - \lim_{n\to\infty} e^{\mu_m^2 n}\mathbb{P}(N_m \geq n+1) \\
&\leq e^{\mu_m^2 (K-1)} \sum_{k=1}^{K-1} \mathbb{P}(N_m = k) + \sum_{k=K+1}^{\infty} \mathbb{P}(N_m \geq k)\left[e^{\mu_m^2 k} - e^{\mu_m^2 (k-1)}\right] + e^{\mu_m^2 K}\mathbb{P}(N_m \geq K) \\
&\leq e^{\mu_m^2 (K-1)} + (1 - e^{-\mu_m^2}) \sum_{k=K}^{\infty} \mathbb{P}(N_m \geq k) e^{\mu_m^2 k} \\
&\leq e^{\mu_m^2 (K-1)} + (1 - e^{-\mu_m^2}) B_m \sum_{k=K}^{\infty} 2^{\frac{-k}{B_m}} e^{\mu_m^2 k} \\
&= e^{\mu_m^2 (K-1)} + (1 - e^{-\mu_m^2}) B_m \frac{(2^{\frac{-1}{B_m}} e^{\mu_m^2})^K}{1 - 2^{\frac{-1}{B_m}} e^{\mu_m^2}} \\
&\leq e^{\mu_m^2 (K-1)} + \mu_m^2 B_m \frac{(2^{\frac{-1}{B_m}} e^{\mu_m^2})^K}{1 - 2^{\frac{-1}{B_m}} e^{\mu_m^2}}
\end{aligned}$$

To use the geometric series formula for the 6th step, we needed $r_m := 2^{\frac{-1}{B_m}} e^{\mu_m^2} < 1$, which is true for m large enough since $\mu_m^2 B_m \to 0$;
additionally, the last step follows from the following inequality: $e^x \leq x + 1$.
Then we use the following fact (remember, $\mu_m^2 B_m \to 0$ by assumption):

$$\liminf_{m\to\infty}(-B_m \log r_m) = \liminf_{m\to\infty}(\log 2 - B_m \mu_m^2) = \log 2$$

This implies that for m big enough, we have

$$\begin{aligned}
-B_m \log r_m \geq \frac{\log 2}{2} &\Leftrightarrow \log r_m \leq \frac{-B_m \log 2}{2} \qquad (2.2) \\
&\Leftrightarrow \frac{1}{1 - r_m} \leq \frac{1}{1 - e^{-\log 2/2B_m}}
\end{aligned}$$

We want to bound this result, and for that we briefly study the following function :

$$\forall x \geq 1, \quad f(x) := \frac{1}{1 - e^{1/x}} \qquad f'(x) = \frac{e^{1/x}}{(x(1 - e^{1/x}))^2}$$

Since

$$\lim_{x\to\infty}\frac{d}{dx}x(1-e^{1/x})=\lim_{x\to\infty}1-e^{1/x}(1-\frac{1}{x})=0$$

we have that the denominator of f' is bounded, and so is f', thus f is Lipschitz continuous. By taking $x=\frac{2B_m}{\log 2}$, we have that for some constant c_1,

$$\frac{1}{1-r_m}\leq\frac{1}{1-e^{-\log 2/2B_m}}\leq c_1B_m$$

For m large enough, we have (from (2.2)) $r_m\leq e^{-\log 2/2B_m}$:

$$\begin{aligned}\mathbb{E}e^{\mu_m^2N_m}&\leq e^{\mu_m^2(K-1)}+\mu_m^2B_m\frac{r_m^K}{1-r_m}\\&\leq e^{\mu_m^2(K-1)}+\mu_m^2c_1B_m^2e^{-\log 2/2B_mK}\\&\leq e^{\mu_m^2(K-1)}+\mu_m^2c_1B_m^2e^{-\log 2/2B_mK}\\&\leq e^{\mu_m^2K}+\mu_m^2c_1B_m^2e^{-\log 2/2B_mK}\end{aligned}$$

Thus for $K=K(m):=\frac{2B_m\log B_m}{\log 2}$, for $c_2:=\frac{2}{\log 2}$, we have

$$\begin{aligned}\mathbb{E}e^{\mu_m^2N_m}&\leq e^{\mu_m^2K}+c_1\mu_m^2B_m^2e^{-K\log 2/2B_m}\\&=e^{\mu_m^2\frac{2B_m\log B_m}{\log 2}}+c_1\mu_m^2B_m^2e^{-\frac{2B_m\log B_m}{\log 2}\frac{\log 2}{2B_m}}\\&=e^{c_2\mu_m^2B_m\log B_m}+c_1\mu_m^2B_m\to 1\end{aligned}$$

since by assumption, $\mu_m^2(\log m)^2\log\log m\to 0$ and B_m is proportional to $(\log m)^2$ by definition.

Since $\mathbb{E}e^{\mu_m^2N_m}\geq 1$, we have $\mathbb{E}e^{\mu_m^2N_m}\to 1$ which was the desired result. $\square$

Chapter 3

Proof theorem 1.2: Regular lattice, bayesian risk

3.1 Part (a): Powerful test

Since by assumption $\mu_m m^{1/4} \to \infty$, we can take a sequence $(h_m)_{m\geq 1}$ of reel numbers growing sufficiently slow, so that $h_m^{-1/2}\mu_m m^{1/4} \to \infty$, and $h_m \to \infty$;
For instance, one could take $h_m := \mu_m m^{1/4}$, but the explicit definition does not actually matter.

Remark 3.1: If we suppose that we do not have $\mu_m \to \infty$ (as the theorem would be true in that case), the first property implies $\frac{\sqrt{m}}{h_m} \to \infty$

The idea is to only consider the nodes at a distance less than $h_m\sqrt{m}$ of the i-axis.
We therefore define the following:
$S(h_m) := \{(i,j) \in V_m, |j| \leq h_m\sqrt{m}\}$, as well as $n_m := |S(h_m)|$

The statistic we will consider is $T_m = \sum_{\nu\in S(h_m)} X_\nu$, and the test we use is to reject the null when $T_m > m\mu_m/2$.

We compute the type I error: Under H_0, T_m has distribution $\mathcal{N}(0, n_m)$, thus (for Z a standard normal random variable)

$$\mathbb{P}_0(T_m > \frac{\mu_m m}{2}) = \mathbb{P}_0(\sqrt{n_m}Z > \frac{\mu_m m}{2}) = \mathbb{P}_0(Z > \frac{\mu_m m}{2\sqrt{n_m}})$$

For this probability to go to 0, we only need $\frac{\mu_m m}{\sqrt{n_m}} \to \infty \Leftrightarrow \frac{\sqrt{n_m}}{\mu_m m} \to 0$
To get there, we have to compute the number n_m of nodes in $S(h_m)$.
We can count those nodes row by row: there is $\lceil (m)/2 \rceil$ of them on the row 0, then $\lceil (m)/2 \rceil + 1$ of them on the row 1, and on and on until the row $\lfloor h_m m \rfloor$. Considering

also the rows below 0, we have

$$\begin{aligned}
n_m &= \lceil (m)/2 \rceil + 2 \sum_{k=1}^{\lfloor h_m\sqrt{m} \rfloor} \lceil (m)/2 \rceil + k \\
&= (\lfloor h_m\sqrt{m} \rfloor + 1)(\lceil (m)/2 \rceil + \frac{\lfloor h_m\sqrt{m} \rfloor}{2}) \\
&\sim (h_m\sqrt{m} + 1)(\frac{h_m\sqrt{m} + m}{2}) \\
&\sim h_m^2 m + m^{3/2} h_m \\
&= h_m m^{3/2}(1 + \frac{h_m}{\sqrt{m}}) \\
&\sim h_m m^{3/2}
\end{aligned}$$

We used the remark (3.1) for the last step. So

$$\frac{\sqrt{n_m}}{\mu_m m} \sim \frac{\sqrt{h_m^2 m}}{\mu_m m} \to \infty$$

This was the last step needed to prove that the type I error goes to 0.
To end the proof, it rests to show that the type II error goes to 0 as well.
We define the event $A_m := \{p \in P_m, \forall \nu \in p, \nu \in S(h_m)\}$ (on the discrete probability space given by P_m).
Again, for Z the standard normal, we have the following:

$$\begin{aligned}
\mathbb{P}_\pi(T_m \leq \frac{\mu_m m}{2}) &= 2^{1-m} \sum_{p \in P_m} \mathbb{P}_{1,p}(T_m \leq \frac{\mu_m m}{2}) \\
&= 2^{1-m} \left[\sum_{p \in A_m} \mathbb{P}_{1,p}(T_m \leq \frac{\mu_m m}{2}) + \sum_{p \notin A_m} \mathbb{P}_{1,p}(T_m \leq \frac{\mu_m m}{2}) \right] \\
&\leq 2^{1-m} \sum_{p \in A_m} \mathbb{P}(\sqrt{n_m} Z + \mu_m m \leq \frac{\mu_m m}{2}) + \pi(A_m^c) \\
&\leq \mathbb{P}(Z \leq \frac{-\mu_m m}{2\sqrt{n_m}}) + \pi(A_m^c)
\end{aligned}$$

As before, the left term goes to 0 since $\frac{\mu_m m}{\sqrt{n_m}} \to \infty$, so we need to prove $\pi(A_m^c) \to 0$ to conclude.
We can do that by considering the paths p as symmetric random walks $(i, S_i)_{i=0}^{m-1}$ with the coordinate i playing the role of "time" or "step".
The of S_i expectation is 0, and the variance is i, since a random walk is a sum of i independent uniform random variables with value in $\{-1, 1\}$.
As a sum of *iid* random variables, $(S_i)_i$ is a martingale, so we can use the Doob's (sub)martingale inequality (for the 2nd step):

$$\pi(A_m^c) = \pi(\sup_{i \leq m-1} |S_i| \geq h_m\sqrt{m}) \leq 2\frac{\mathbb{E}[|S_{m-1}|]}{h_m\sqrt{m}} \leq 2\frac{\sqrt{\mathbb{E}[S_{m-1}^2]}}{h_m\sqrt{m}} = 2\frac{\sqrt{m}}{h_m\sqrt{m}} = \frac{2}{h_m} \to 0$$

Since both Type I and Type II error go to 0, the risk (i.e. their sum) goes to 0 as well, so our test is powerful. □

3.2 Part (b): All tests are powerless

We now suppose that $\mu_m \log(m)^{-1/4} \to 0$
We have

$$B_m(\pi) := \inf_{testT} \gamma_\pi(T) = \mathbb{P}_0(L_m \geq 1) + \mathbb{P}_\pi(L_m < 1)$$

We also have

$$B_m(\pi) = 1 - \frac{\mathbb{E}_0|L_m - 1|}{2} \geq 1 - \frac{\sqrt{\mathbb{E}_0(L_m - 1)^2}}{2}$$

as in Theorem 1.1 (b).
This inequality implies that if we can show $\mathbb{E}_0(L_m - 1)^2 \to 0$, then any test is asymptotically powerless and we are done.
By section (1.4), we have

$$\mathbb{E}_0(L_m - 1)^2 = \mathbb{E}e^{\mu_m^2 N_m} - 1$$

with $N_m : (P_m^2, 2^{P_m^2}, \pi^2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ the number of crossings between two paths.
Now, we need to study the random variable N_m in order to obtain a useful bound (from above).
As we have seen in the proof of theorem 1.1 (b), we can consider a path as a random walk $S = (i, S_i)_{i=0}^m$ with the coordinate i playing the role of "time". With this notation, we have to find the (random) number N_m of indexes such that $S_i = S'_i$, for two independent random walks S and S'; equivalently, we want for the difference $S - S'$to be equal to 0. Note that since $S' = -S'$ in distribution, we might as well consider the sum.
For each step, each random walk can either go up or go down of 1, and we will then sum up both values; we notice that this is exactly the outcome we get by considering two steps of only one random walk, i.e. staying put with probability 1/2 and going up or down with probability 1/4.
So finally, we arrive to the conclusion that N_m is simply the number of return to 0 of the random walk $(S_{2i})_{i=1}^m$, where we go two steps at a time, with the index going up to $2m$.
From here, we use a result from ([5], page 96) which gives us precisely the following:

$$\mathbb{P}(N_m = k) = \frac{1}{2^{2m-k}} \binom{2m-k}{m} = \frac{1}{2^{2m-k}} \frac{(2m-k)!}{m!(m-k)!}$$

What is left to do is to bound the right-hand side from above, using (the upper side of) Stirling's approximation formula:

$$n! < 2\pi^{n+\frac{1}{2}} e^{-n+\frac{1}{12n}}$$

Using it, we have for $k < m$

$$\begin{aligned}\mathbb{P}(N_m = k) &\le (\pi m)^{-\frac{1}{2}} \frac{(1-\frac{k}{2m})^{2m-k+\frac{1}{2}}}{(1-\frac{k}{m})^{m-k+\frac{1}{2}}} \\ &= (\pi m)^{-\frac{1}{2}} \sqrt{\frac{(1-\frac{k}{2m})}{(1-\frac{k}{m})}} e^{-mg(\frac{k}{m})}\end{aligned}$$

where we define $g(x) := (1-x)\log(1-x) - 2(1-\frac{x}{2})\log(1-\frac{x}{2})$ and we have $g'(x) = \log(1-\frac{x}{2}) - \log(1-x)$, $g''(x) = \frac{1}{x-2} + \frac{1}{x-1}$

For $x \in (0,1)$, since $g''(x) > 1/2$, we have $g(x) > -\frac{t^2}{4}$ for $x \in (0,1)$; we will use this result later for the computations.

Fact: for $x \in \left[0, \frac{1}{2}\right]$, we have $\frac{1-x/2}{1-x} \le 1 + x$ (since it is equivalent to $1 - \frac{x}{2} \le 1 - x^2$)

Now, we consider $\epsilon \in \left[0, \frac{1}{2}\right]$. We will use two different bounds, depending on whether $k \le \epsilon m$: If $k \le \epsilon m$, we have $\frac{1-k/2m}{1-k/m} \le 1 + k/m \le 1 + \epsilon$ by the previous fact.

If $\epsilon m < k < m$, we use the more general bound $\sqrt{\frac{1-k/2m}{m(1-k/m)}} \le 1$, which follows directly from the inequality $k \le m$. If $k = m$, we simply have $\mathbb{P}(N_m = m) = 2^{-m}$

With these results, we can bound the interesting quantity:

$$\begin{aligned}\mathbb{E}e^{\mu_m^2 N_m} &= \sum_{k=1}^{m} e^{\mu_m^2 k}\mathbb{P}(N_m = k) \\ &\le 2^{-m} + \sum_{k=1}^{m-1} e^{\mu_m^2 k}(\pi m)^{-\frac{1}{2}} \sqrt{\frac{(1-\frac{k}{2m})}{(1-\frac{k}{m})}} e^{-mg(\frac{k}{m})} \\ &\le 2^{-m} + \frac{1}{\sqrt{\pi}} \sum_{k=1}^{m-1} e^{\mu_m^2 k} \sqrt{\frac{(1-\frac{k}{2m})}{m(1-\frac{k}{m})}} e^{-\frac{k^2}{4m}} \\ &\le 2^{-m} + \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\lfloor \epsilon m \rfloor} e^{\mu_m^2 k - \frac{k^2}{4m}} \sqrt{\frac{1+\epsilon}{m}} + \frac{1}{\sqrt{\pi}} \sum_{k=\lfloor \epsilon m \rfloor + 1}^{m-1} e^{\mu_m^2 k} e^{-\frac{k^2}{4m}}\end{aligned}$$

We can see that the second sum goes to 0:

$$\sum_{k=\lfloor \epsilon m \rfloor + 1}^{m-1} e^{\mu_m^2 k} e^{-\frac{k^2}{4m}} \le m e^{\mu_m^2 (m-1)} e^{-\frac{\epsilon^2 m}{4}} = m e^{m(\mu_m^2 - \epsilon^2) - \mu_m^2}$$

The right hand-side goes to 0 since $\mu_m^2 - \frac{e^2}{4} \le \delta$ for some $\delta < 0$ and m large enough. We focus now on the first sum.

Because $\mu_m^2 k - \frac{k^2}{4m}$ and thus $e^{\mu_m^2 k - \frac{k^2}{4m}}$ is (strictly) decreasing in k for $k \ge \frac{\mu_m^2}{2m}$, we can bound the sum by an integral for $k \ge 1 \ge \frac{\mu_m^2}{2m}$:

$$\begin{aligned}
\frac{1}{\sqrt{\pi m}}\sum_{k=0}^{\lfloor \epsilon m\rfloor} e^{\mu_m^2 k-\frac{k^2}{4m}} &\leq \frac{1}{\sqrt{\pi m}}+\frac{1}{\sqrt{\pi m}}\int_0^{\epsilon m} e^{\mu_m^2 t-\frac{t^2}{4m}}dt \\
&= \frac{1}{\sqrt{\pi m}}+\sqrt{\frac{2}{\pi}}\int_0^{\epsilon\sqrt{m/2}} e^{\sqrt{2m}\mu_m^2 x-\frac{x^2}{2}}dx \qquad\qquad t=\sqrt{2m}x \\
&= \frac{1}{\sqrt{\pi m}}+2\frac{e^{\mu_m^4 m}}{\sqrt{2\pi}}\int_0^{\epsilon\sqrt{m/2}} e^{-\frac{1}{2}(x-\mu_m^2\sqrt{2m})^2}dx \\
&= \frac{1}{\sqrt{\pi m}}+2e^{\mu_m^4 m}\mathbb{P}(Z\in\left[\mu_m^2\sqrt{2m},(\mu_m^2+\epsilon)\sqrt{2m}\right]) \\
&\leq \frac{1}{\sqrt{\pi m}}+2e^{\mu_m^4 m}\mathbb{P}(Z\geq\mu_m^2\sqrt{2m})
\end{aligned}$$

Since by assumption $\mu_m m^{1/4}\to 0$, the result above goes to 1. Since $\mathbb{E}e^{\mu_m^2 N_m}\geq 1$, we have

$$\lim_{m\to\infty}\mathbb{E}e^{\mu_m^2 N_m}=1$$

We can now put together all the pieces and conclude:

$$\liminf_{m\to\infty} B_m(\pi)\geq \liminf_{m\to\infty} 1-\frac{\sqrt{\mathbb{E}_0(L_m-1)^2}}{2}=\lim_{m\to\infty} 1-\frac{\sqrt{\mathbb{E}e^{\mu_m^2 N_m}-1}}{2}=1$$

Thus any sequence of test is powerless. $\square$

Chapter 4

Proof theorem 1.3: Binary tree

For the following proof, we will use the generalized likelihood ratio test (GLRT). It is based on the following statistic Λ (the generalized likelyhood ratio), that we compute:

$$\begin{aligned}\Lambda(X_\nu, \nu \in V_m) :&= \frac{\sup_{p \in P_m} f_{X,H_{1,p}}(X_\nu, \nu \in V_m)}{f_{X,H_0}(X_\nu, \nu \in V_m)} \\ &= \sup_{p \in P_m} \frac{\prod_{\nu \in p} f_{X_\nu,H_{1,p}}(X_\nu)}{\prod_{\nu \in V} f_{X_\nu,H_0}(X_\nu)} \\ &= \sup_{p \in P_m} \exp(\sum_{\nu \in p} X_\nu \mu_m - \frac{\mu_m^2}{2}) \\ &= \exp(\sup_{p \in P_m} X_p \mu_m - \frac{\mu_m^2}{2})\end{aligned}$$

Since the GLRT is given by comparing this statistic to a threshold, and that Λ is a strictly increasing function of $\sup_{p \in P_m} X_p$, we can focus only on the latter.
It explains the choice of the statistic M_m for the following proof.

4.1 Part (a): Powerful test

We will consider the cases $\mu > \sqrt{2 \log 2}$ and $\mu = \sqrt{2 \log 2}$ separately, starting with the former (we follow the process of the original paper).
The test we use in that case is rejecting the null if
$M_m := \sup\{X_p, p \in P_m\}$ is strictly bigger than $m\sqrt{2 \log 2}$
We first prove

$$\mathbb{P}_0(M_m \geq m\sqrt{2 \log 2}) \to 0$$

We use the following bound for the tail of a standard gaussian Z:

$$\mathbb{P}(Z \geq t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx \leq \frac{1}{\sqrt{2\pi}} \int_t^\infty \frac{x}{t} e^{-x^2/2} dx = \frac{e^{-t^2/2}}{t\sqrt{2\pi}}$$

which gives us for any $p \in P_m$, recalling that under the null X_p is distributed as

$\mathcal{N}(0, m)$:

$$\begin{aligned}
\mathbb{P}_0(M_m \geq m\sqrt{2\log 2}) &= \mathbb{P}_0(X_p \geq \sqrt{2m\log 2}, \forall p \in P_m,) \\
&\leq 2^{m-1}\mathbb{P}_0(X_p \geq m\sqrt{2\log 2}) \qquad \text{for some } p \in P_m \\
&= 2^{m-1}\mathbb{P}(Z \geq \sqrt{2m\log 2}) \\
&\leq \frac{1}{4\sqrt{\pi m \log 2}} \qquad (4.1) \\
&\to 0
\end{aligned}$$

Now, for an hypothesis $H_{1,p}$ with $p \in P_m$, we have

$$\frac{M_m}{m} \geq \frac{X_p}{m} \sim \mathcal{N}(\mu, \frac{1}{m}) \to \mu > \sqrt{2\log 2}$$

(with convergence in probability) With that, we can conclude that if $\mu > \sqrt{2\log 2}$, then for any $p \in P_m$, $\mathbb{P}_{1,p}(M_m > m\sqrt{2\log 2}) \to 1$ The case $\mu > \sqrt{2\log 2}$ is therefore finished since both probability of Type I and Type II errors go to 0.

Now, if $\mu = \sqrt{2\log 2}$, we cannot use anymore the fact that $\frac{M_m}{m}$ stays above $\sqrt{2\log 2}$ for m large enough, since it could not be the case. We will instead use the following result for the case of H_1:

$$\liminf_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} \mathbb{I}\{\frac{M_{m^k}}{m^k} \geq \mu\} \geq 1/2 \qquad (4.2)$$

with $m_k := 2^k$.
To prove inequality (4.2), notice that since $M_m \geq X_p(m)$ for p the elevated path, it suffices to show

$$\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} \mathbb{I}\{X_p(m_k) \geq \mu m_k\} = 1/2 \qquad (4.3)$$

To see this, notice that $X_p(m)$ is a sum of m *iid* random variables distributed as $\mathcal{N}(\mu, 1)$ If we take $N_j, j \geq 1$ independent standard gaussians, and $Z_i := \sum_{j=1}^{i} N_j$, (4.3) is equivalent to

$$\lim_{k\to\infty} \frac{1}{k}\sum_{i=1}^{k} \mathbb{I}\{Z_{2^i} \geq 0\} = 1/2$$

which follows from which follows from Ergodic Theorem for the fast mixing random walks Z_i.
(for the rest of the proof, we take limits as $k \to \infty$ instead of m) The test we consider is to claim H_0 if and only if $S_k := \sum_{i=1}^{k} \mathbb{I}\{M_{m_i} > m_i\sqrt{2\log 2}\} < \frac{k}{4}$
i.e. , we count the number of times $M_{m_i} \geq m_i\sqrt{2\log 2}$ and we compare it with $\frac{k}{4}$.
From (4.2), we have that for any path $p \in P_m$, we have $\mathbb{P}_{1,p}(S_k \geq \frac{k}{4}) \to 1$, or equivalently $\mathbb{P}_{1,p}(S_k < \frac{k}{4}) \to 0$.
Since it works for any path p, we also have $\sup_{p\in P_m} \mathbb{P}_{1,p}(S_k < \frac{k}{4}) \to 0$.

Finally, for the Type I error, we use the same upper bound (4.1) as before, this time with m_k instead of m, to get

$$\sum_{k\geq 1} \mathbb{P}_0(M_{m_k} \geq m_k\sqrt{2\log 2}) \leq \sum_{k\geq 1} \frac{1}{4\sqrt{\pi m_k \log 2}} = \frac{1}{4\sqrt{\pi \log 2}} \sum_{k\geq 1} \frac{1}{2^{k/2}} < \infty$$

Using the Borel-Cantelli lemma, we have that

$$\mathbb{P}_0(M_{m_k} \geq m_k\sqrt{2\log 2} \text{ infinitely often }) = 0$$

Hence, the number of times $M_{m_i} \geq m_i\sqrt{2\log 2}$ is bounded with ($\mathbb{P}_0$)-probability 1, which implies that the Type I error goes to 0 (see the definition of the test).

Since the Type II error also goes to 0, we have $\gamma(\mathbb{I}\{S_k < \frac{k}{4}\}) \to 0$ □

4.2 Part (b): No powerful test

Suppose now that $\mu < \sqrt{2\log 2}$
We use again the Likelihood Ratio

$$L_m = 2^{-(m-1)} \sum_{p\in P_m} \exp(X_p\mu - \frac{m\mu^2}{2})$$

By proposition 1 in [6], L_m is a (non-negative) martingale with respect to the filtration $\mathcal{F}_m := \sigma(X_\nu, |\nu| \leq m)$.
The martingale convergence theorem implies that L_m converges almost surely to a finite non-negative random variable L_∞.
We will use the following result:

$$\begin{aligned} B_m(\pi) &= \mathbb{P}_0(L_m \geq 1) + \mathbb{P}_\pi(L_m < 1) \\ &= 1 - \mathbb{P}_0(L_m \geq 1) + \mathbb{P}_\pi(L_m \leq 1) \\ &= 1 - \mathbb{E}_0(\mathbb{I}_{\{L_m \leq 1\}}) + \mathbb{E}_0(L_m\mathbb{I}_{\{L_m\leq 1\}}) \\ &= 1 - \mathbb{E}_0\left[(1-L_m)_+\right] \end{aligned}$$

By dominated convergence, we have the following:

$$\lim_{m\to\infty} B_m(\pi) = 1 - \mathbb{E}_0\left[(1-L_\infty)_+\right] \qquad (4.4)$$

Using Proposition 2 in [6], we have that L_m is uniformly integrable.
This is equivalent (for a martingale) to L^1 convergence, thus we deduce $\mathbb{E}_0[L_\infty] = \lim_{m\to\infty} \mathbb{E}_0[L_m] = 1$. Thanks to this, we now have (remembering that L_∞ in non-negative) $\mathbb{P}_0(L_\infty = 0) < 1$, which in turn implies

$$l := \lim_{m\to\infty} B_m(\pi) > 0$$

by equation (4.4).
This implies that for any test, its bayesian risk is bounded away from 0 and thus such a test cannot be powerful.
Since by definition the minimax risk is greater (or equal) than the bayesian risk, the conclusion holds for both.
□

4.3 Part (c): All tests are powerless

We now assume that $\mu_m \to 0$.
We proceed as in Theorem 1.1 (b); we simply have to show $Var_0(L_m) = \mathbb{E}e^{\mu_m N_m} - 1 \to 0$, where N_m is the number of cross between two paths.
In the current case, N_m is actually very easy to compute. For two valid paths is the binary tree, we notice that if two paths do not share a vertex, then all the vertices further from the root will not be shared as well. If we construct the path from the root, then for each step the two path will branch apart with probability 1/2. The event $N_m = k$ happens if both paths did not branch before the k-th vertex, but branch at the $(k+1)$-th; we can therefore deduce for $1 \leq k < m$:

$$\mathbb{P}(N_m = k) = 2^{-k} \text{ and } \mathbb{P}(N_m = m) = 2^{-m+1}$$

(for k=m, the probability is twice greater since the path ends and thus cannot branch after the m-th vertex) Now, we can compute:

$$\begin{aligned}
\mathbb{E}e^{\mu_m N_m} &= \sum_{k=1}^{m} e^{\mu_m^2 k} 2^{-k} + e^{\mu_m^2 m} 2^{-m} \\
&= \sum_{k=0}^{m} (\frac{e^{\mu_m^2}}{2})^k + (\frac{e^{\mu_m^2}}{2})^m - 1 \\
&= \frac{1 - (\frac{e^{\mu_m^2}}{2})^{m+1}}{1 - \frac{e^{\mu_m^2}}{2}} + (\frac{e^{\mu_m^2}}{2})^m - 1
\end{aligned}$$

From $\mu_m \to 0$ we get that $(\frac{e^{\mu_m^2}}{2})^m \to 0$, which implies that the result above goes to 1. Finally, we get $Var_0(L_m) = \mathbb{E}e^{\mu_m N_m} - 1 \to 0$, so we can conclude that any test is powerless (with respect to the bayesian risk, and hence also with respect to the minimax risk.). □

Chapter 5

Conclusion

We have been able to show for different setups and for two different type of graphs that the detection of an elevated path was either possible or that it was not, depending on the assumptions we make about the decaying rate of the elevation; there are however several questions that are left open.

The first is to know whether it would be possible to cover all the possible decaying rates, for completeness; we have seen that in some cases, the Theorem 1.1 could not conclude anything, namely when the decaying rate is in between $\sqrt{\log x}$ and $\log x\sqrt{\log\log x}$. The article from Arias-Castro et al. also leaves this question open.

Another would be to generalize the problem to other graphs. Such graphs should however come in different sizes, to make it is possible to consider the limit as the size goes to infinity. The original article generalizes the results to the regular lattice in arbitrary dimension, but we could thing of other such graphs: complete k-ary tree, regular lattice going in all directions, or even maybe randomly generated graphs, to name a few. The set of valid paths considered could also change, to allow for regressions or discontinuities for example.

Finally, another generalization angle is the one of the random variables considered. The article from Arias-Castro generalizes to the exponential family, but there still other possibilities, such as the uniform distribution for instance.