DYNAMIC OVERSHOOT CORRECTION FOR SPRT TESTS

1. The bus example

Imagine one is informed that every hour two buses arrive at a certain stop. One probably expects an average waiting time of 15 minutes when arriving at a random moment. However this is not true if the second bus always arrives 10 minutes after the first one. In that case the expected waiting time is

$$(50/60) \times (50/2) + (10/60) \times (10/2) = 2600/120 = 21.66 \text{ minutes}$$

which is considerably more than 15 minutes.

2. Discussion of the bus example

Let us generalize the example by assuming that N buses arrive at the stop with intervals $D_1, \ldots, D_N$. It we arrive at a random time the probability of catching bus i is $D_i / \sum_j D_j$ and in that case the expected waiting time is $D_i/2$. In other words the amount of time we expect to wait is

$$\frac{\sum_i D_i^2}{2 \sum_i D_i} \tag{2.1}$$

If the D_i are iid copies of a random variable D then the expected waiting time can be estimated as

$$\frac{E(D^2)}{2E(D)} \tag{2.2}$$

Moreover (2.1) can be used to estimate the expected waiting time by simply observing the arrival of buses for some time.

3. Application to random walks

Consider a random walk

$$S_i = L_1 + L_2 + \cdots + L_i$$

where the L_i are iid. The random walk finishes when $(S_i)_i$ leaves a given interval $[A, B]$.

Assume that $(S_i)_i$ leaves the interval through the right boundary. I.e. for some τ we have

$$\begin{aligned} S_i &< B \qquad \text{for } i < \tau \\ S_\tau &\geq B \end{aligned}$$

We are interested in the expected value of $S_\tau - B$.

We write

$$S_\tau = L_1 + \cdots + L_{j_1} + L_{j_1+1} + \cdots + L_{j_2} + \cdots + L_{j_{t-1}+1} + \cdots + L_{j_t}$$

where

$$D_k := L_{j_{k-1}+1} + \cdots + L_{j_k} \tag{3.1}$$

satisfies $D_k \geq 0$ and $L_{j_{k-1}+1} + \cdots + L_{j_{k-1}+u} < 0$ for $u < j_k - j_{k-1}$. Note that the probability distribution of D_k has to be conditioned on the existence of an L_{j_k} such that the righthand side of (3.1) is ≥ 0. If $E(L_i) < 0$ this is not automatic.

The $(L_i)_i$ are iid random variables, so that the D_k are also iid. Let the D_k be copies of D.

If B is large compared to $|S_i|$ then it is reasonable to assume that the "bus model" applies so that the expected values of $S_\tau - B$ is given by (2.2). Moreover the value of (2.2) can be estimated by observing the $(D_i)_i$ themselves during the random walk via the the formula (2.1).

A similar discussion applies when the random walk leaves the interval $[A, B]$ via the left boundary.

4. Application to SPRT tests

To be written.