(1) For this problem pretend we are dealing with a language with six-word dictionary

\{\text{fun, sun, sit, sat, fan, for}\}.

An extensive study of literature written in this language reveals that all words are equally likely except that “for” is \(\alpha\) times as likely as the other words. Further study reveals that

(a) Each keystroke is an error with probability \(\theta\)
(b) All letters are equally likely to produce errors
(c) Given that a letter is typed incorrectly it is equally likely to be any other letter
(d) Errors are independent across letters.

For example, the probability of correctly typing “fun” (or any other word) is \((1 - \theta)^3\), the probability of typing “pun” or “fon” when intending to type is “fun” is \(\theta(1 - \theta)^2\), and the probability of typing “foo” or “nnn” when intending to type “fun” is \(\theta^2(1 - \theta)\).

Use Bayes rule to develop a simple spell checker for this language. For each of the typed words “sun”, “the”, “foo”, give the probability that each word in the dictionary was the intended word, e.g., given they typed “the” what is the probability they were trying to type “fun”? Perform this for (i) \(\alpha = 2\) and \(\theta = 0.1\); (ii) \(\alpha = 50\) and \(\theta = 0.1\); and (iii) \(\alpha = 2\) and \(\theta = 0.95\). Briefly comment on the changes you observe in these three cases.

(2) A steel plant relies on a machine that produces devices guaranteed to be defecting with probability less than 0.10. You are in charge of quality control. To ensure that the machine is working properly you take 100 samples each day (independent across sample and day) and record the number of samples that are defective. The data for a recent 10-day stretch are given below:

<table>
<thead>
<tr>
<th>Day</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of defective samples</td>
<td>12</td>
<td>11</td>
<td>9</td>
<td>15</td>
<td>18</td>
<td>15</td>
<td>11</td>
<td>8</td>
<td>12</td>
<td>16</td>
</tr>
</tbody>
</table>

Assuming the true defect rate is constant within this relatively short period of time, plot the posterior probability that the defect probability is greater than 0.1 as a function of the day. For day \(t\), use the samples from day 1 through day \(t\) in this calculation. When, if ever, would you sound the alarm and claim that the machine is out of order?

You should turn in your responses to these questions in 1-2 pages (i.e., one piece of paper with text on both sides). Be sure all plots are labeled and code is commented!