Part 1

You are consulting for a hospital. They have a diagnostic test for a disease with a known background **prevalence of 0.1%.**

The test has the following properties:

**p(positive result | person has disease) = 0.91**

**p(negative result| person does not have disease) = 0.84**

The cost of running the test one time is $1. The test can be repeated for each patient and the results of the test are independent of one another allowing for Bayesian updates. The test always yields a positive or negative result.

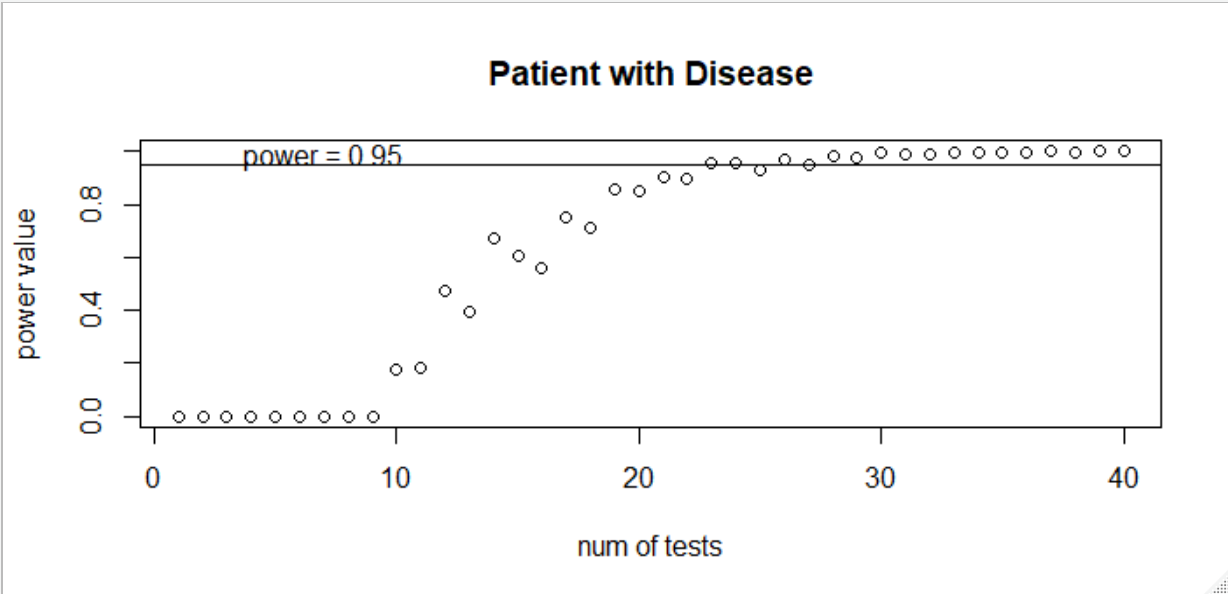
The requirement of the hospital is that the test is repeated for each patient until a Bayesian posterior

of at least **0.9999 is reached for >=95% of patients**.

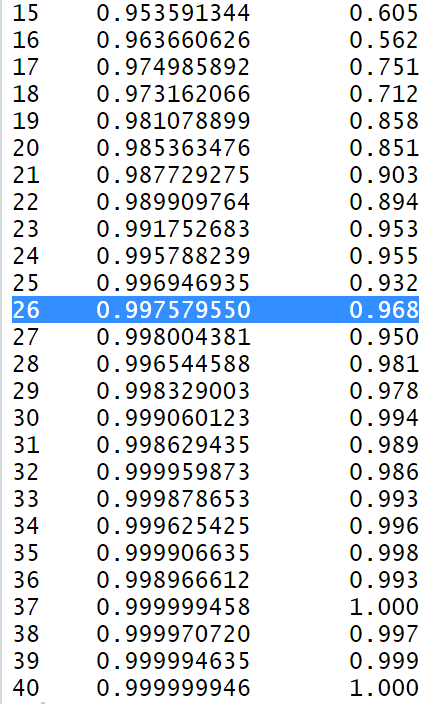
1. Run simulations for a patient with the disease. About how many times on average must the test be repeated to achieve the hospital’s requirements?

The test must be repeated about 26 times for a patient with the disease (figures 1A & 1B).

* [bayes\_with\_disease\_part\_1.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/bayes_with_disease_part_1.R)
* [run\_bayes\_sim.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/run_bayes_sim.R)



**Figure 1A (Power Values per # of Tests for Patient with Disease):** P(Test+ | Disease+) = 0.91;   
π = (Disease+ = 0.001, Disease-- = 0.999 )

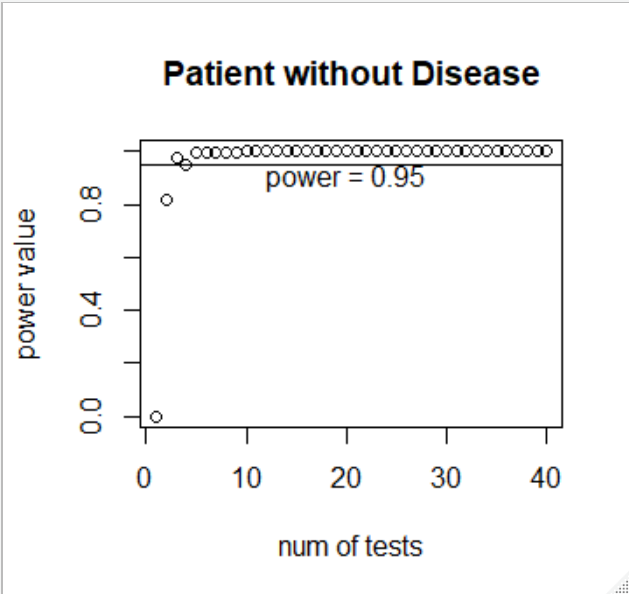


**Figure 1B (Power Values per # of Tests for Patient with Disease):** This figure shows that it took about 26 tests to meet the hospital’s requirements (data for figure 1A).

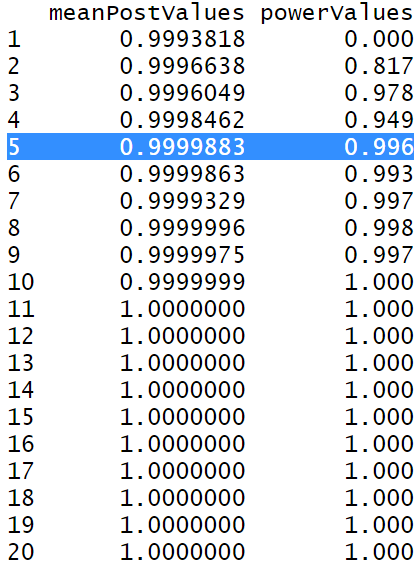
1. Repeat the simulations for a patient without the disease. About how many times on average must the test be repeated to achieve the hospital’s requirements?

The test must be repeated about five times for a patient without the disease (figures 2A & 2B).

* [bayes\_without\_disease\_part\_1.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/bayes_without_disease_part_1.R)
* [run\_bayes\_sim.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/run_bayes_sim.R)



**Figure 2A (Power Values per # of Tests for Patient without Disease):** P(Test-- | Disease--) = 0.84;   
π = (Disease+ = 0.001, Disease-- = 0.999 )



**Figure 2B (Power Values per # of Tests for Patient without Disease):** This figure shows that it took about 5 tests to meet the hospital’s requirements (data for figure 2A).

(3) The hospital plans to run the test on one million patients per year. At a cost of $1 per test, about how much should the hospital budget to run these tests? (That is to say, for a million patients, how many tests can the hospital anticipate running?)

Show your work/code/justification for all answers…

ASSUMPTION:

I’m assuming the hospital has a policy to focus on the trustworthiness of true positive tests [P(Test+ | Disease+)]. That is, if a patient tests positive for the disease, the hospital wants to ensure the diagnosis is highly accurate. Therefore, the hospital should budget $26M (26 test repeats \* 1M patients \* $1).

**Part 2**

Another manufacturer approaches the hospital with an improved, but more expensive, test with the following properties

**p(positive result | person has disease) = 0.96**

**p(negative result| person does not have disease) = 0.95**

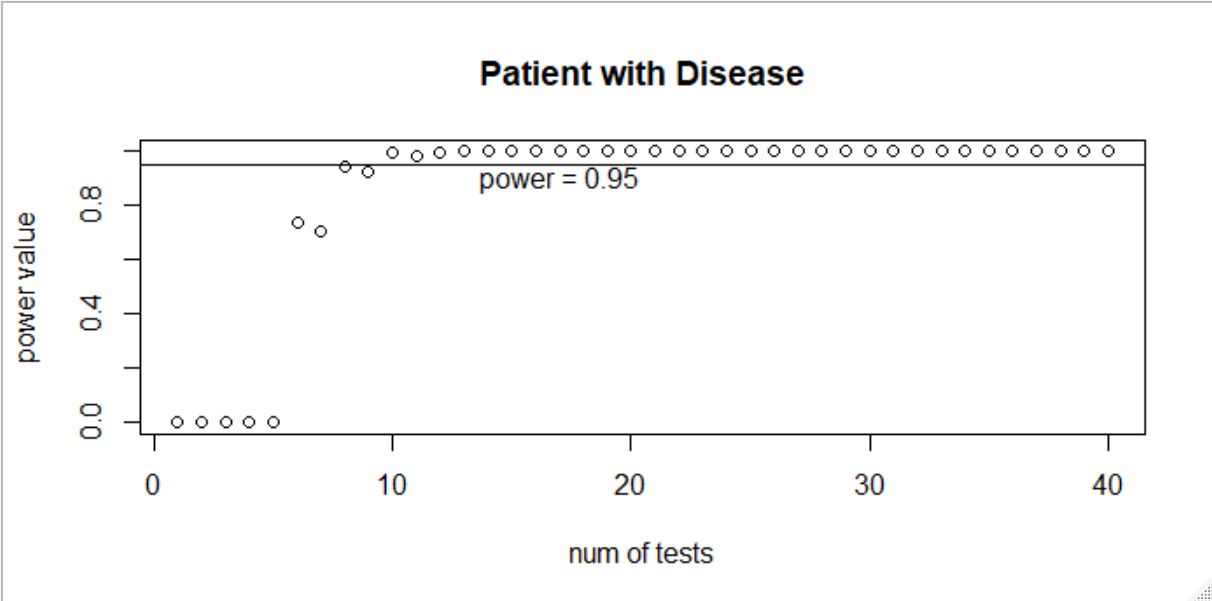
(1) With this test, how many times on average must the test be repeated to achieve the hospital’s requirements for patients with and without the disease?

The test must be repeated about 10 times for a patient with the disease (figures 3A & 3B).

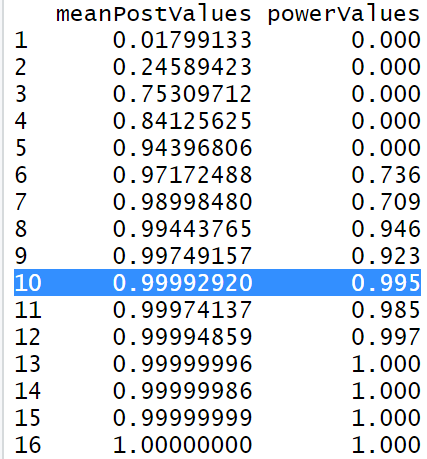
* [bayes\_with\_disease\_part\_2.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/bayes_with_disease_part_2.R)
* [run\_bayes\_sim.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/run_bayes_sim.R)

The test must be repeated about three times for a patient without the disease (figures 4A & 4B).

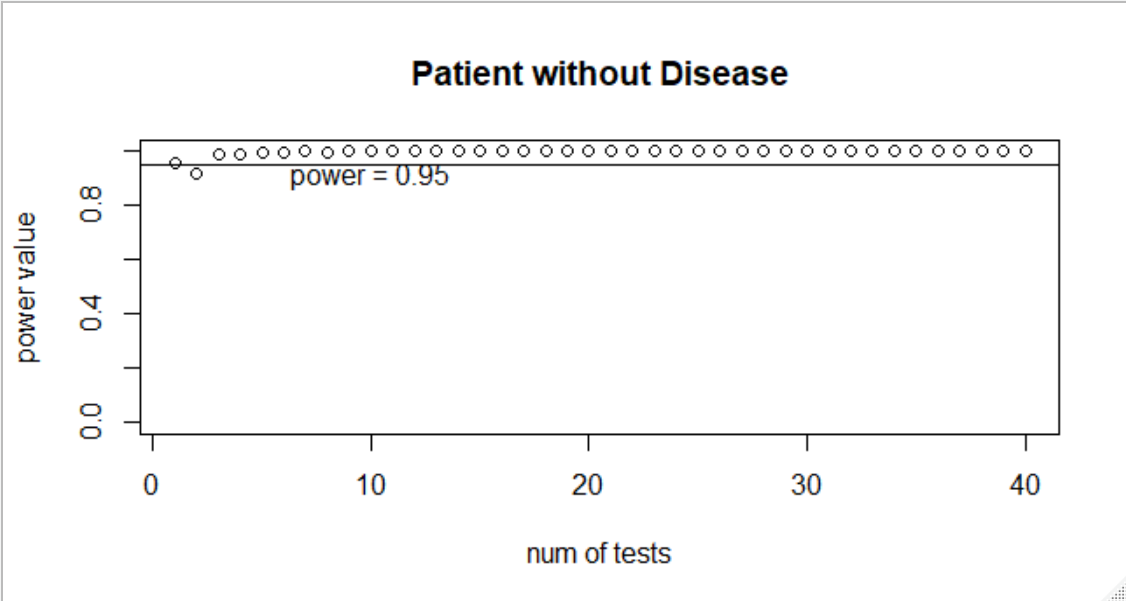
* [bayes\_without\_disease\_part\_2.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/bayes_without_disease_part_2.R)
* [run\_bayes\_sim.R](https://github.com/jyoung67/advstatistics-labs/blob/master/labs/lab03/run_bayes_sim.R)



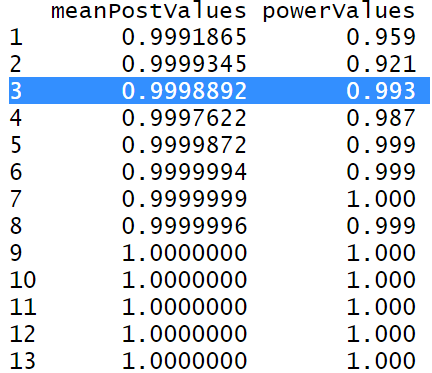
**Figure 3A (Power Values per # of Tests for Patient with Disease):** P(Test+ | Disease+) = 0.96;   
π = (Disease+ = 0.001, Disease-- = 0.999 )



**Figure 3B (Power Values per # of Tests for Patient with Disease):** This figure shows that it took about 10 tests to meet the hospital’s requirements (data for figure 3A).



**Figure 4A (Power Values per # of Tests for Patient without Disease):** P(Test-- | Disease--) = 0.95;   
π = (Disease+ = 0.001, Disease-- = 0.999 )



**Figure 4B (Power Values per # of Tests for Patient with Disease):** This figure shows that it took about three tests to meet the hospital’s requirements (data for figure 4A).

(2) Considering only the cost of the test, and assuming the hospital will screen one million patients with a background prevalence of 0.1%, at about what price point for running the test one time will the hospital save money by switching to the new test?

In order for the hospital to save money by switching to the new test, the cost for a single test would need to be less than $2.60 (see calculation below).

R1\*P1\*β = σ (current test)

R2\*P2\* β = σ =>

P2 = (R1\*P1)/R2 = (26\*$1)/10 = $2.60

Note: R is # of repeats per test; P is price per single test; β is expected # of patients per year; σ is total budget