Advanced Algorithms 2016. Exercise 9

Exercise 9

This exercise has a twofold purpose:
(1) practicing proper calculations with probabilities in an algorithmic context, and
(2) discouraging knowledge gaps in exams.

A student takes a course with n topics. For simplicity we consider only binary outcomes for each topic: The student has either learned the topic very well, or has not learned the topic at all.
Specifically, suppose that the student has learned $n-k$ topics and has not learned k topics. The examiner applies an extremely simple randomized algorithm: (S)he chooses t times a random topic and asks a question about it. Each of the n topics is chosen with the same probability, and the t choices are independent. The same topic may be chosen repeatedly (perhaps with different questions). For the highest mark, all topics must be learned. Let p be the probability that no knowledge gap is detected, that is, all randomly selected topics have been learned, and the student earns the highest mark.

(a) How large is p (exactly or approximately), as a function of n, k, t?
You may assume that n and t are large numbers, and k is small compared to n, such that you can take advantage of the approximate equation $(1-x)^{1/x} \approx 1/e$ for small x, where $e = 2.718\ldots$. This yields a nicer expression for the result. Explain how you calculated your result. In particular, point out where the assumption of independent choices is essential.

(b) Discuss: What can happen if we drop the independence assumption?

(c) How does p behave if the examiner doubles the number t of questions?

(d) Depending on n and k, how many questions does the examiner need to detect a knowledge gap with probability $1/2$?
For simplicity you can give the number t in O-notation, that is, only the rough size of t, ignoring constant factors.