1. To save time and money when testing blood samples for the presence of a disease, the following procedure is sometimes used: Blood samples from a number of people are pooled together and the pooled sample is tested. If the pooled sample is negative, then no further testing need be done. If the pooled sample is positive, however, then each of the separate samples must be tested individually. To be specific, suppose that the blood samples of 10 people are pooled together in this procedure, and suppose that the proportion of those who have the disease in the population from which the samples are drawn is .02.

(a) What is the probability that only one test must be run?
(b) What is the probability that eleven tests must be run?
(c) Let \( Y \) be the number of tests that must be run. Use your answers to (a) and (b) to write down the probability distribution of \( Y \).
(d) Do you think this procedure is more efficient than just testing the 10 samples separately?

Solutions.
(a) Let \( X \) represent the number of people of the 10 who have the disease. Then \( X \) has a binomial distribution with \( n = 10 \) and \( \pi = 0.02 \). The only way that only one test must be run is if none of the people have the disease, i.e., if \( X = 0 \). The probability of this is

\[
P(X = 0) = \frac{10!}{0!10!} (0.02)^0 (0.98)^{10} \approx 0.817.
\]

(b) The only possible numbers of tests are 1 and 11. Since we know that the probability of 0 tests is 0.817, we know that the probability of 11 tests is \( 1 - 0.817 = 0.183 \).
(c) \( P(Y = 1) = 0.817 \); \( P(Y = 11) = 0.183 \).
(d) Yes, since there’s a large chance (0.817) that we’ll only need one test, and a small chance (0.183) that we’ll need 11 tests. (We could formally compute the mean of \( Y \) to be \( (1)(0.817) + (11)(0.183) = 2.83 \) and note that this is smaller than 10.)

2. Answer the above problem in the case that the population proportion of those who have the disease is .1. Is your answer to part (d) the same?

Solutions.
We use the same technique, except we replace \( \pi = 0.02 \) by \( \pi = 0.1 \). (a) Let \( X \) represent the number of people of the 10 who have the disease. Then \( X \) has a binomial distribution with \( n = 10 \) and \( \pi = 0.1 \). The only way that only one test must be run is if none of the people have the disease, i.e., if \( X = 0 \). The probability of this is

\[
P(X = 0) = \frac{10!}{0!10!} (0.1)^0 (0.9)^{10} \approx 0.349.
\]

(b) The only possible numbers of tests are 1 and 11. Since we know that the probability of 0 tests is 0.349, we know that the probability of 11 tests is \( 1 - 0.349 = 0.651 \).
(c) $P(Y = 1) = 0.349$; $P(Y = 11) = 0.651$.
(d) It’s a little harder to say. There’s a greater than 50% chance that we’ll need 11 tests, but there’s still a pretty high (0.349) chance that we’ll need only one test. (We could formally compute the mean of $Y$ to be $(1)(0.349) + (11)(0.651) = 7.51$ and note that this is smaller than 10.)

3. Answer the above problem in the case that the population proportion of those who have the disease is .3. Is your answer to part (d) the same?

Solutions.
We use the same technique, except we replace $\pi = 0.1$ by $\pi = 0.3$. (a) Let $X$ represent the number of people of the 10 who have the disease. Then $X$ has a binomial distribution with $n = 10$ and $\pi = 0.3$. The only way that only one test must be run is if none of the people have the disease, i.e., if $X = 0$. The probability of this is

$$P(X = 0) = \frac{10!}{0!10!} (0.3)^0 (0.7)^{10} \approx 0.028.$$  

(b) The only possible numbers of tests are 1 and 11. Since we know that the probability of 0 tests is 0.028, we know that the probability of 11 tests is $1 - 0.028 = 0.972$.
(c) $P(Y = 1) = 0.028$; $P(Y = 11) = 0.972$.
(d) No. There’s a very high (0.972) chance that we’ll need 11 tests, and a very small (0.028) chance that we’ll need only one test. (We could formally compute the mean of $Y$ to be $(1)(0.028) + (11)(0.972) = 10.72$ and note that this is larger than 10.)