Expected number of solutions for random CNF formulas

This is to clear up some confusion about the calculation of the expected number of solutions of conjunctive normal form (CNF) formulas in the August 25 lecture. I apologize that my notation and language failed to communicate an idea that is actually quite simple. Here's a second try.

Let $x_1, \ldots, x_m$ be the Boolean variables of the CNF formula. Consider a formula with just two clauses, y_1 and y_2. Both clauses are selected at random from a set Y, and the probability of selecting $y \in Y$ is $p(y)$. For example, we might have (for $m = 4$)

$$Y = \{x_1 \vee \bar{x}_2 \vee x_4,\ x_1 \vee x_3 \vee x_4,\ \bar{x}_2 \vee \bar{x}_3 \vee x_4\},$$

that is, just three possibilities with positive probability. In the random ensemble discussed in lecture, Y included all $\binom{m}{3} \times 2^3$ ways of combining three distinct variables in a disjunction with negations and their probabilities $p(y)$ were equal. The factorization property does not depend on the choice of $p(y)$

The 2-clause CNF $z = y_1 \wedge y_2$ is drawn with probability $p(y_1)p(y_2)$. This is the only assumption we need to get factorization of our expectation value – that the clause-choice random variables are independent and drawn from the same distribution. Now suppose we have some assignment, say $x = (\mathrm{F}, \mathrm{T}, \mathrm{T}, \mathrm{F})$. The probability that $z = \mathrm{T}$ for this assignment, in our random formula ensemble, is then

$$\mathrm{prob}(z = \mathrm{T}) = \sum_{y_1 \in Y} \sum_{y_2 \in Y} p(y_1)p(y_2)\ \delta(y_1(x) = \mathrm{T})\ \delta(y_2(x) = \mathrm{T}),$$

where $y_1(x)$ is the truth value of clause y_1 when given assignment x, etc. The Kroenecker delta's implement the conjunction, since only when each one equals 1 is their product 1 and $z = \mathrm{T}$. As you can see, the double sum factorizes:

$$\mathrm{prob}(z = \mathrm{T}) = \left(\sum_{y \in Y} p(y)\ \delta(y(x) = \mathrm{T}) \right)^2 .$$

The sum in parentheses is the probability that an individual clause is true for the assignment x, and depends on x. In lecture we made use of the additional property that this probability is independent of x (and equal to 7/8) when the random clause ensemble has the uniform distribution. The expected number of solutions is then the product of the number of x's (truth assignments) and the probability of $z = \mathrm{T}$.