Axis-aligned rectangles:

I actually found the explanation in the textbook (as well as every other explanation I could find online) to be quite unclear, so I am going to briefly re-derive it here using a different approach that I find much clearer than any of the material I could find online -- I couldn’t find a single “proof” in any lecture notes or textbooks online that did not have a major flaw in it (not that the equations were necessarily wrong, but just that the logical flow was broken in several places).

We are looking for the sufficient conditions under which the error of our axis-aligned rectangles learning will be less than some This condition is as follows: that there exist four rectangles along the sides of R, {r­1, r­2, r­3, r­4} such that , and all four sides of our model Rs are within these rectangles. Now since we are only looking for the sufficient condition, we do not need to worry about all possible sets of rectangles {r­1, r­2, r­3, r­4}, but instead can simply pick four whose areas are all equal to Note that this does not necessarily give the tightest possible bound, and it is very possible that by optimizing the areas of the four side rectangles, we could get a tighter bound.

The probability of a side being inside rectangle ri is the probability that at least one positive point lie inside that rectangle, or equivalently that not all positive points lie outside that rectangle, i.e. all outside ri], and then the probability that this is simultaneously true for all four rectangles is . By DeMorgans law, this probability is equal to . The probability that all m points lie outside rectangle ri is . Therefore, by the union bound, the sufficient condition for R(Rs)< occurs with probability .