8.

Intro

In this question, we have to prove the equivalence between two definitions of conditional independence. The alternative definition says that we can decouple the probability function into two functions with arguments (X,Z) and (Y,Z) iff we have conditional independence. The intuition behind the alternative definition is based on the chain rule: $p(x, y|z) = p(x|z)p(y|x, z)$. Note that the first term is just a function of X and Z (g(x,z)). On the other hand, the second term is a function of the three r.v's: X, Y and Z. If $X \perp Y|Z$ then we can decouple the functions, as we wanted. Furthermore, if we know that $p(x, y|z)$ can be decoupled, then it makes sense that $p(y|x, z) = p(y|z) = h(y, z)$. Now that we have some intuition on the problem, let's go to the actual proof.

Solution

This is a if and only if question.

Let's start with the if part:

$p(x, y|z) = g(x, z)h(y, z) \implies p(x, y|z) = p(x|z)p(y|z)$

First, Let's sum all possibles cases of x and y on the pdf. We are assuming discrete r.v's. For continuous r.v's just change the summation by an integral.

$$\sum_x \sum_y p(x, y|z) = 1 \tag{1}$$

At the same time:

$$\begin{aligned} \sum_x \sum_y p(x, y|z) = \sum_x \sum_y g(x, z)h(y, z) = \\ (\sum_x g(x, z))(\sum_y h(y, z)) \end{aligned} \tag{2}$$

From (1) and (2) we arrive at:

$$(\sum_x g(x, z))(\sum_y h(y, z)) = 1 \tag{3}$$

Now, let's calculate the marginal distributions of $p(x, y|z)$

$$\begin{aligned} p(y|z) = \sum_x p(x, y|z) = \sum_x g(x, z)h(y, z) = \\ h(y, z) \sum_x g(x, z) \end{aligned} \tag{4}$$

Isolating, $\sum_x g(x, z)$, we get:

$$\sum_x g(x, z) = \frac{p(y|z)}{h(y, z)} \tag{5}$$

$$p(x|z) = \sum_y p(x,y|z) = \sum_y g(x,z)h(y,z) = \\ g(x,z)\sum_y h(y,z) \tag{6}$$

Isolating, $\sum_y h(y,z)$, we get:

$$\sum_y h(y,z) = \frac{p(x|z)}{g(x,z)} \tag{7}$$

Finally, multipling (5) and (7) and substituting (3) on the result we get:

$$(\sum_x g(x,z))(\sum_y h(y,z)) = \frac{p(y|z)}{h(y,z)}\frac{p(x|z)}{g(x,z)} = 1 \tag{8}$$

Thus, $h(y,z) = p(y|z)$ and $g(x,z) = p(x|z)$ and we proved that $p(x,y|z) = g(x,z)h(y,z) \implies p(x,y|z) = p(x|z)p(y|z)$.

Now, let's prove the only if part:

$p(x,y|z) = p(x|z)p(y|z). \implies p(x,y|z) = g(x,z)h(y,z)$

This is much more easier to prove. $p(x|z)$ is only a function of x and z, so $p(x|z) = g(x,z)$. Using the same argument for y and z, we get $p(y|z) = h(y,z)$. So, $p(x,y|z) = p(x|z)p(y|z). \implies p(x,y|z) = g(x,z)h(y,z)$.

Conclusion In this exercise we proved one alternative way of defining conditional independance. This result is more important that it may look at first glance. Initially, we know that conditional independance was achieved if we could decouple the **conditional probabilities** of X and Y. With this result, we know that the decomposition of $p(x,y|z)$ into **any** pair of functions g(x,z) and h(y,z) implies conditional probability. As we all know is much more easy to look to a function and conclude 'look, it only have X/Y and Z terms' that to prove that each function is in fact a probability function.