A correct proof of Lemma 5.1 from InfoGAN

Nic Ford*
Avital Oliver†

May 2018

The InfoGAN paper[1] has the following lemma:

Lemma 5.1. *For random variables X, Y and function $f(x, y)$ under suitable regularity conditions:* $\mathbb{E}_{x\sim X, y\sim Y|x}[f(x,y)] = \mathbb{E}_{x\sim X, y\sim Y|x, x'\sim X|y}[f(x,y)]$.

The proof in the paper seems wrong – here's a step where x mysteriously becomes x':

$$\int_x \int_y P(x,y) f(x,y) \int_{x'} P(x'|y) dx' dy dx$$
$$= \int_x P(x) \int_y P(y|x) \int_{x'} P(x'|y) f(x', y) dx' dy dx$$

After consulting with others, we couldn't fix this proof. Instead, Nic Ford found the following proof:

Proof.

$\mathbb{E}_{x\sim X, y\sim Y\|x}[f(x,y)] =$	make expectations explicit...
$\mathbb{E}_{x\sim P(X)}\left[\mathbb{E}_{y\sim P(Y\|X=x)}[f(x,y)]\right] =$	by definition of $P(Y\|X=x)$...
$\mathbb{E}_{x,y\sim P(X,Y)}[f(x,y)] =$	by definition of $P(X\|Y=y)$... ...
$\mathbb{E}_{y\sim P(Y)}\left[\mathbb{E}_{x\sim P(X\|Y=y)}[f(x,y)]\right] =$	rename x to x'...
$\mathbb{E}_{y\sim P(Y)}\left[\mathbb{E}_{x'\sim P(X\|Y=y)}[f(x',y)]\right] =$	by the law of total expectation...
$\mathbb{E}_{x\sim P(X)}\left[\mathbb{E}_{y\sim P(Y\|X=x)}\left[\mathbb{E}_{x'\sim P(X\|Y=y)}[f(x',y)]\right]\right] =$	make expectations implicit...
$\mathbb{E}_{x\sim X, y\sim Y\|x, x'\sim X\|y}[f(x',y)]$	

□

*http://nicf.net
†http://aoliver.org
[1] https://arxiv.org/pdf/1606.03657.pdf