STATISTICAL RETHINKING WINTER 2020/2021
HOMEWORK, WEEK 4

1. Consider three fictional Polynesian islands. On each there is a Royal Ornithologist charged by the king with surveying the birb population. They have each found the following proportions of 5 important birb species:

	Birb A	Birb B	Birb C	Birb D	Birb E
Island 1	0.2	0.2	0.2	0.2	0.2
Island 2	0.8	0.1	0.05	0.025	0.025
Island 3	0.05	0.15	0.7	0.05	0.05

Notice that each row sums to 1, all the birbs. This problem has two parts. It is not computationally complicated. But it is conceptually tricky.

First, compute the entropy of each island's birb distribution. Interpret these entropy values.

Second, use each island's birb distribution to predict the other two. This means to compute the K-L Divergence of each island from the others, treating each island as if it were a statistical model of the other islands. You should end up with 6 different K-L Divergence values. Which island predicts the others best? Why?

2. Recall the marriage, age, and happiness collider bias example from Chapter 6. Run models `m6.9` and `m6.10` again. Compare these two models using WAIC (or LOO, they will produce identical results). Which model is expected to make better predictions? Which model provides the correct causal inference about the influence of age on happiness? Can you explain why the answers to these two questions disagree?

3. Reconsider the urban fox analysis from last week's homework. Use WAIC or LOO based model comparison on five different models, each using `weight` as the outcome, and containing these sets of predictor variables:

(1) `avgfood + groupsize + area`
(2) `avgfood + groupsize`
(3) `groupsize + area`
(4) `avgfood`
(5) `area`

Can you explain the relative differences in WAIC scores, using the fox DAG from last week's homework? Be sure to pay attention to the standard error of the score differences (`dSE`).