Probability Kinematics and Halpern's Full Employment Theory

Stefan Lukits

May 22, 2011

1. Introduction

Let f be a probability distribution on a finite space $x_1, \ldots, x_m$ that fulfills the constraint

$$\sum_{i=1}^{m} r(x_i) f(x_i) = \alpha \tag{1}$$

Because f is a probability distribution it fulfills

$$\sum_{i=1}^{m} f(x_i) = 1 \tag{2}$$

We want to maximize the entropy, given the constraints (1) and (2),

$$-\sum_{i=1}^{m} f(x_i) \ln(x_i) \tag{3}$$

We use Lagrange multipliers to define the functional

$$J(f(x_i)) = -\sum_{i=1}^{m} f(x_i) \ln f(x_i) + \lambda_0 \sum_{i=1}^{m} f + \lambda_1 \sum_{i=1}^{m} r(x_1) f(x_i) \tag{4}$$

and differentiate it with respect to x_i

$$\frac{\partial J}{\partial f(x_i)} = -\ln(f(x_i)) - 1 + \lambda_0 + \lambda_1 r(x_i) \tag{5}$$

Set (5) to 0 to find the necessary condition to maximize (3)

$$g(x_i) = e^{\lambda_0 - 1 + \lambda_1 r(x_1)} \tag{6}$$

This is the Gibbs distribution. We still need to do two things: (a) show that the entropy of g is maximal, and (b) show how to find λ_0 and λ_1. (a) is shown in Theorem 12.1.1 in Cover and Thomas [1] and there is no reason to copy it here. I was not able to find (b) in the literature and will show it here. A faint suggestion on how to do it is in Jaynes' treatment of the Brandeis Dice Problem (see [2, 243]).

Let

$$\lambda_1 = -\beta \tag{7}$$

$$Z(\beta) = \sum_{i=1}^{m} e^{-\beta r(x_i)} \tag{8}$$

$$\lambda_0 = 1 - Z(\beta) \tag{9}$$

To find λ_0 and λ_1 we introduce the constraint

$$-\frac{\partial}{\partial \beta} \ln(Z(\beta)) = \alpha \tag{10}$$

To see how this constraint gives us λ_0 and λ_1 Jaynes' solution of the Brandeis Dice Problem is a helpful example. We are, however, interested in a general proof that this choice of λ_0 and λ_1 gives us the probability distribution maximizing the entropy. That g so defined maximizes the entropy is shown in (a). We need to make sure, however, that with this choice of λ_0 and λ_1 the constraints (1) and (2) are also fulfilled.

First, we show

$$\sum_{i=1}^{m} g(x_i) = \sum_{i=1}^{m} e^{\lambda_0 - 1 + \lambda_1 r(x_i)} = e^{\lambda_0} \sum_{i=1}^{m} e^{\lambda_1 r(x_i)} = \\ e^{-\ln(Z(\beta))} Z(\beta) = 1 \tag{11}$$

Then, we show, by differentiating $\ln(Z(\beta))$ using the substitution $x = e^{-\beta}$

$$\alpha = -\frac{\partial}{\partial \beta} \ln(Z(\beta)) = -\frac{1}{\sum_{i=1}^{m} x^{r(x_i)}} \left(\sum_{i=1}^{m} r(x_i) x^{r(x_i) - 1} \right) (-x) = \\ \frac{\sum_{i=1}^{m} r(x_i) x^{r(x_i)}}{\sum_{i=1}^{m} x^{r(x_i)}} \tag{12}$$

And, finally,

$$\sum_{i=1}^{m} r(x_i)g(x_i) = \sum_{i=1}^{m} r(x_i)e^{\lambda_0 - 1 + \lambda_1 r(x_1)} = e^{\lambda_0 - 1}\sum_{i=1}^{m} r(x_i)e^{\lambda_1 r(x_1)} =$$

$$e^{\lambda_0 - 1}\sum_{i=1}^{m} r(x_i)x^{r(x_i)} = \alpha e^{\lambda_0 - 1}\sum_{i=1}^{m} x^{r(x_i)} = \alpha e^{\lambda_0 - 1}\sum_{i=1}^{m} e^{-\beta r(x_i)} =$$

$$\alpha Z(\beta)e^{\lambda_0 - 1} = \alpha Z(\beta))e^{-\ln(Z(\beta))} = \alpha \tag{13}$$

References

[1] Cover, T. and Thomas, J., 2006. *Elements of Information Theory*, volume 6. Wiley, Hoboken, NJ.

[2] Jaynes, E., 1989. *Papers on Probability, Statistics and Statistical Physics.* Springer, Dordrecht.