Lifted Threshold Variable Models

Erin Bennett

Contents

RSA: Inferred Threshold

Math

Literal Listener's probability distribution over the values X are is prior, conditioned on the utterance being true and renormalized.

$$P_{L0}(x|u,\theta) \propto \delta_{u \text{ is true}} \cdot P(x)$$

Speaker's utility is the negative cost and the log probability of the actual state of the world under the Literal Listener's posterior. This means that the more surprised the Literal Listener would be to hear the true state of the world after already hearing the utterance, the less good the utterance would be.

$$\mathbb{U}_S(u|x,\theta) = log(P_{L0}(x|u,\theta)) - cost(u)$$

The speaker then chooses an utterance by soft-maximizing their utility funciton.

$$P_S(u|x,\theta) \propto e^{\lambda \mathbb{U}_S(u|x,\theta)}$$

The pragmatic listener infers both the threshold θ and the value x conditioning on the speaker choosing the given utterance.

$$P_{L1}(x,\theta|u) \propto P_S(u|x,\theta)P(x)P(\theta)$$

Code (WebPPL)

```
var literal_listener = cache(function(utterance, theta) {
  Enumerate(function() {
        var value = value_prior();
        factor(meaning(utterance, theta, value) ? 0 : -Infinity);
        return value;
    })
```

```
})

var speaker = cache(function(value, theta) {
    Enumerate(function() {
        var utterance = utterance_prior();
        var literal_interpretation = literal_listener(utterance, theta);
        factor(literal_interpretation.score([], value) * speaker_lambda);
        return utterance;
    })
})

var listener = function(utterance) {
    var value = value_prior();
    var theta = theta_prior();
    var speaker_choice = speaker(value, theta);
    factor(speaker_choice.score([], utterance));
    return [value, theta];
}
```

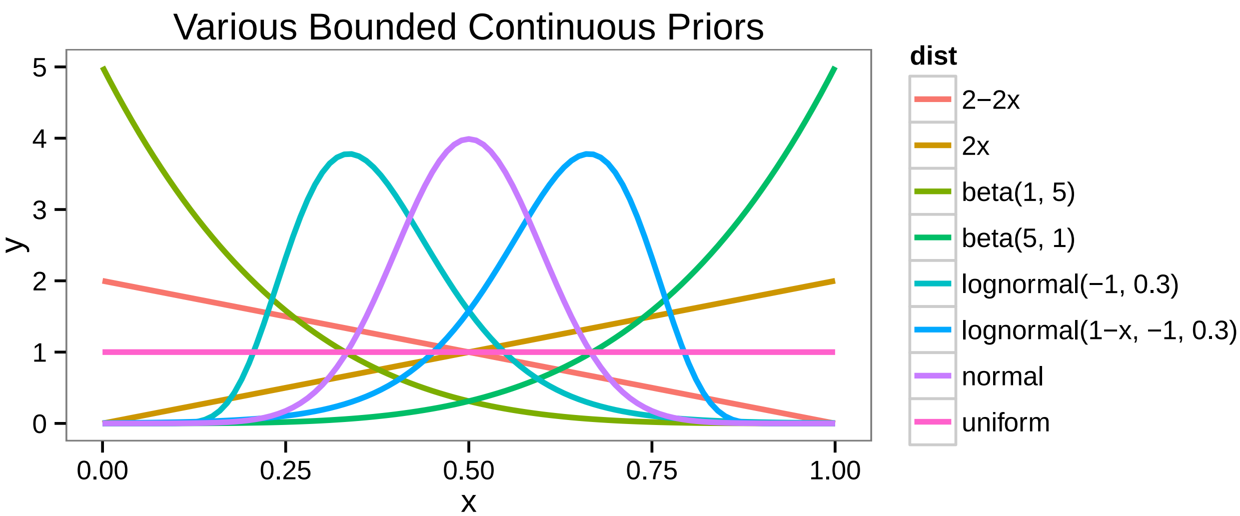
Model Simulations for Some Theoretical Prior Distributions

Priors

I looked at several different prior distributions on the bounded interval [0,1]:

- (bounded) normal: $\delta_{[0,1]} \cdot \mathcal{N}(0.5, 0.1)$
- uniform: $\mathcal{U}(0, 1)$
- betas:
- $\text{Beta}(1, 5)$
- $\text{Beta}(5, 1)$
- (bounded) log-normal:
- $\delta_{[0,1]} \cdot \ln \mathcal{N}(-1, 0.3)$
- reversed: $\delta_{[0,1]} \cdot \ln \mathcal{N}(1 - x; -1, 0.3)$
- linear:
- $2 - 2x$
- $2x$

These distributions were discretized for the simulations.



Simulation Results