Modeling Everyday Predictions

We talked about how the Bayesian model proposed by Griffiths and Tenenbaum (2006) can predict human judgments on some everyday prediction problems.

In this tutorial, we are going to implement their model, with human lifespan as an example. **LifespanExercise.m** is a MATLAB worksheet. It guides you through how to implement the model in a step-by-step manner. Below is the framework:

1. Define the **input value** `t` (the current age of the person, given in the problem). Let's make t = 80

2. Define a vector `ttotal`, which covers the **range of all possible estimates** (hypotheses)

3. Compute the **prior probability** for each value in `ttotal`:

$ttotal \sim N(\mu, \sigma^2)$, where μ = mean human lifespan (78), σ = standard deviation (13)
* use normpdf function in matlab for normal prob distribution function
* Normalize the prior probabilities (i.e., make them sum to 1)

4. Compute the **likelihood** for observing `t`, based on values in `ttotal`:

$$P(t|ttotal) = \begin{cases} \frac{1}{ttotal}, & t \geq ttotal \\ 0, & 0 < t < total \end{cases}$$

5. Apply the Bayes rule to compute the **posterior probability** for each value in `ttotal`, given `t`:

$P(tt\text{otal} \,|t) \propto P(t|\, tt\text{otal})P(tt\text{otal})$
* Normalize the posterior probabilities (i.e., make them sum to 1)

6. Obtain the **median** of the posterior as answer. In other words, we want to locate the value t_{median} such that:

Σ posterior probs to the LEFT of $t_{median} \approx \Sigma$ posterior probs to the RIGHT of t_{median}
$\Rightarrow \Sigma$ posterior probs to the LEFT of $t_{median} \approx 0.5$

Σ probs to the LEFT of xx is called the **cumulative probability** at x.
Therefore, our goal is to locate t_{median} whose cumulative probability is equal (or closest) to 0.5.

Follow the steps below to find t_{median}:
i. Obtain the cumulative probabilities on the posterior using the `cumsum()` function: If A = [2 8 3 5 1], `cumsum(A)` returns [2 2+8 2+8+3 2+8+3+5 2+8+3+5+1]. The cumulative probabilities should be the cumulative sums of Posterior×stepsize.
ii. Subtract 0.5 from each value in the cumulative probability vector obtained in i).
iii. Compute the absolute value for each value in the difference vector obtained in ii).
iv. Find the position kk in the vector obtained in iii) which gives you the minimum value.
v. Obtain the median of posterior by retrieving the k^{th} value in `ttotal`.

Hint: From time to time, you may want to visualize the probabilities in prior, likelihood, and posterior:
```
plot(ttotal, TheProbVectorYouWantToPlot)
```

[In-class exercise, q1.m]

Now you’ve constructed the model that is capable of estimating humans’ judgment for any input t. Let’s obtain a plot of the estimates for a range of input t!

1. Define a vector `T`, storing a range of possible input values `t` (e.g., from 1:100).

2. Write a FOR loop to obtain an estimate from your model for each value in `T`. Below are the pseudocode:

> *For* each value `T(i)`
>> Use the code we have already written to get an estimate for the lifespan at T(i)
>> Store the answer in the ith position in the vector `ModelPred`
>
> *`End`*`;`

3. Plot `ModelPred` (as Y) against `T` (as X)