Winter 2019/20 TERM 1 EPSE 581C: ASSIGNMENT 1
Due: Monday, Oct. 7th

- Please make sure you write your answers to these questions in your own words. Even if you work with a group to formulate your responses, do not just copy someone else's sentences/words.
- There is no need to record more than 2 decimal places for any of these problems; HOWEVER, do not round numbers until you obtain a final answer.
- All problem data are available online in .csv format.

Question 1: (Updating beliefs with new information via Bayes' Theorem) Suppose we randomly select a single Vancouver resident from the municipal voting registry (their identity has been hidden for legal privacy reasons!).

(a) What is a reasonable prior probability that the sampled individual is female?

(b) Suppose the sampled individual voted Green in the last provincial election. Based on previous voting data, it is known that about 11% of females in Vancouver voted Green, while about 7% of males did so. Using this information and your answer from part (a), find the (updated) probability that the sampled individual is female using Bayes' Theorem.

(c) Suppose that the sampled individual also worked as a voting clerk during the last provincial election (an important civic duty!). It is known that females are twice as likely to volunteer their time as election staffers than are males (this is true!), and this tendency is independent of which party one votes for. Using this information, update the probability that the sampled individual is female using Bayes' Theorem and your answer from part (b).

(d) Finally, suppose that the sampled individual is known to hold a freshwater fishing license. It is known that 1 female holds a fishing license for about every 3 males that do in BC (this is true!). Moreover, this tendency holds regardless of which party one votes for or whether one volunteers to help with elections. Use this last bit of information (along with all the previous information) to update the probability that the sampled individual is female.

(e) Would your answer to part (d) change if you had changed the order of parts (b) and (c)? That is, does the Bayesian updating procedure depend on the order in which you include new information? Why or why not?

Question 2: (Joint, marginal, and conditional mass functions) Let X and Y be two independent Bernoulli(0.5) random variables and define $U = X + Y$ and $V = X - Y$.

(a) Find the joint and marginal probability mass functions for U and V. [It is sufficient to construct a table to describe these mass functions.]

(b) Are U and V independent? Why or why not?

(c) Find the conditional probability mass functions $\Pr(U = u \mid V = v)$ for all permissible v and $\Pr(V = v \mid U = u)$ for all permissible u. [Again, you can construct a table to describe these mass functions.]

Question 3: (Bayesian estimation and inference with exponentially distributed phenomena) Suppose we are interested in studying social media habits of people on Twitter with a low number of followers (apparently, the average number of followers is now 707 for a typical Twitter user (!!!)). We randomly select 5 Twitter users, each with about 100 followers. We are interested in quantifying how long it takes for a new tweet to receive its first "like". We observe each of the 5 Twitter users from 12 PM to 2 PM (local time) on a Monday (apparently the busiest time for Twitter), and record how long it takes their first tweet in that time frame to receive its first "like." This random phenomenon is a continuous random variable, and can be modelled reasonably well by an *Exponential* distribution. That is, if $X \sim Exp(\lambda)$ for some $\lambda > 0$, then X has density function

$$f_X(x) = \lambda e^{-\lambda x}, \quad x \geqslant 0.$$

Using some basic calculus, one can show that the mean and the standard deviation of this distribution are both equal to $1/\lambda$.

Suppose we collect the following sample data (time is measured in seconds from when the tweet was posted):

$$\{10, 45, 12, 62, 34\}$$

(a) Assuming these Twitter users and their followers are independent of each other (perhaps they share no common followers), write down the *likelihood* function that describes these observed data, $f(\boldsymbol{y} \mid \lambda)$.

(b) Now we need to specify a prior for λ. Since λ can, theoretically, assume any positive value, we need a prior distribution that is defined over $\mathbb{R}^+$. The Exponential distribution seems like a natural choice. Suppose we decide to use an Exponential prior with mean equal to 10 seconds. Write down the density for this prior distribution, $\pi(\lambda)$.

(c) Using Bayes' Theorem, write down the posterior distribution for the observed data under this assumed prior. Recall that the normalizing factor is just a *constant*. Thus, you may write your answer as

$$f(\lambda \mid \boldsymbol{y}) = c \cdot f(\boldsymbol{y} \mid \lambda)\pi(\lambda).$$

(d) This posterior distribution is an example of a *Gamma* distribution (and Exponentials are special cases of Gammas). In general, Gamma random variables have the following density function:

$$f_W(w) = C \cdot w^{\alpha-1} e^{-\beta w}, \quad w \geqslant 0,$$

where $\alpha, \beta > 0$ are shape parameters (here, β is on the *inverse scale* to λ above). One can show (using basic calculus again) that the mean of this Gamma distribution is β/α (on the appropriate scale). Using this fact, what is the mean of the posterior in part (c)? How does this answer make sense in the context of the data you observed and the prior that we assumed?

(e) Play around with the Gamma (and Exponential) densities with this applet (nothing to report!): `https://homepage.divms.uiowa.edu/~mbognar/applets/gamma.html`