1. Alice is 50% sure that Bob's email address is `352642@acme.com`. She runs a test by sending 100 emails to random recipients of the form $abcdef$`@acme.com`. 95 of the emails bounce as undeliverable. Alice then sends an email intended for Bob to `352642@acme.com` and it does not bounce. What is the (posterior) probability that she correctly guessed his email address? *[Adapted from Bertsekas-Tsitsiklis problem 8.1.1]*

   Solution: The random variable of interest Θ is an indicator for the event that Bob's email address is `352642@acme.com`. Alice's prior PMF is $\mathrm{P}(\Theta = 0) = \mathrm{P}(\Theta = 1) = 1/2$.

   Let X be the indicator for the event that Alice's email to `352642@acme.com` doesn't bounce, then $\mathrm{P}(X = 1|\Theta = 1) = 1$. We can also reasonably assume that $\mathrm{P}(X = 1|\Theta = 0)$ is about 5%. I.e., if Alice is wrong, her email is about 5% likely to land in an actual mailbox.

   The desired posterior probability is

   $$\mathrm{P}(\Theta = 1|X = 1) = \frac{\mathrm{P}(X = 1|\Theta = 1) \cdot \mathrm{P}(\Theta = 1)}{\mathrm{P}(X = 1)} \propto \frac{1}{2},$$

   where the normalization constant is

   $$\mathrm{P}(X = 1) = \mathrm{P}(X = 1|\Theta = 1)\,\mathrm{P}(\Theta = 1) + \mathrm{P}(X = 1|\Theta = 0)\,\mathrm{P}(\Theta = 0) = 1 \cdot \frac{1}{2} + 0.05 \cdot \frac{1}{2}$$

   from where $\mathrm{P}(\Theta = 1|X = 1) = 1 \cdot \frac{1}{2}/(1 \cdot \frac{1}{2} + 0.05 \cdot \frac{1}{2}) = 1/1.05 \approx 0.95$.

2. After much dating experience, Romeo concludes that girls show up Exponential(Θ) hours late to a date, where Θ is a Uniform$(0, 1)$ random variable that describes a girl's type.

   (a) Juliet is 10 minutes late on their first date. What is Romeo's posterior PDF for Θ?

   Solution: Let X_1 be the time of Juliet's arrival in hours.
   The PDF of X_1 conditioned on Θ is $f_{X_1|\Theta}(x_1|\theta) = \theta e^{-\theta x_1}$ (for $x_1 \geq 0$).
   The posterior PDF is

   $$f_{\Theta|X_1}(\theta|1/6) \propto f_{X_1|\Theta}(1/6|\theta) \cdot f_\Theta(\theta) = \theta e^{-\theta/6} \cdot 1 = \theta e^{-\theta/6}$$

   when $0 \leq \theta \leq 1$, and 0 otherwise.
   As the conditional PDF must integrate to one, this actual value must equal to $\theta e^{-\theta/6}$ divided by $\int_0^1 \theta e^{-\theta/6} d\theta$. Using an online integrator (or integration by parts) the integral evaluates to $-(6x + 36)e^{-x/6}|_0^1 = 36 - 42e^{-1/6} \approx 0.448$, so $f_{\Theta|X_1}(\theta|1/6) \approx 2.233\theta e^{-\theta/6}$ for $0 \leq \theta \leq 1$.

   (b) Juliet is 30 minutes late on their second date. Assume Juliet's late time is independent of one another. What is Romeo's posterior PDF now?

   Solution: After the second date,

   $$\begin{aligned} f_{\Theta|X_1,X_2}(\theta|1/6, 1/2) &\propto f_{X_1|\Theta}(1/6|\theta) \cdot f_{X_2|X_1,\Theta}(1/2|\theta) \\ &= f_{X_1|\Theta}(1/6|\theta) \cdot f_{X_2,\Theta}(1/2|\theta) \\ &= \theta e^{-\theta/6} \cdot \theta e^{-\theta/2} \\ &= \theta^2 e^{-2\theta/3}. \end{aligned} \tag{1}$$

   To obtain the value, we normalize by $\int_0^1 \theta^2 e^{-2\theta/3} = (27 - 51e^{-2/3})/4 \approx 0.203$. Therefore, the posterior PDF is $f_{\Theta|X_1,X_2}(\theta|1/6, 1/2) = 4/(27 - 51e^{-2/3}) \cdot \theta^2 e^{-2\theta/3} \approx 4.904\theta^2 e^{-2\theta/3}$ for $0 \leq \theta \leq 1$.

3. The number of mahjong games played in a given family on the Lunar New Year can be modeled as a Geometric(Θ) random variable, where the value of the parameter Θ varies from family to family.

   (a) Assuming a Beta(α, β) prior PDF on Θ, what is its posterior PDF given that x_1 games were played on the last Lunar New Year? (**Hint:** It is also a Beta-type random variable.) What if the prior on Θ is Uniform$(0, 1)$?

   Solution: The conditional PDF of X_i given Θ is Geometric(Θ), namely $f_{X_i|\Theta}(x_i|\theta) = (1-\theta)^{x_i-1}\theta$, so the posterior PDF is

$$\begin{aligned} f_{\Theta|X_1}(\theta|x_1) &\propto f_{X_1|\Theta}(x_1|\theta) \cdot f_\Theta(\theta) \\ &= (1-\theta)^{x_1-1}\theta \cdot (1-\theta)^{\beta-1}\theta^{\alpha-1} \\ &= (1-\theta)^{(x_1+\beta-2)}\theta^{\alpha}, \end{aligned}$$

   which is the PDF of a Beta($\alpha+1, x_1+\beta-1$) random variable. In particular, if the prior is Uniform$(0, 1)$ which is the same as Beta$(1, 1)$ the posterior will be Beta$(2, x_1)$.

   (b) Now suppose that you have more data, i.e. you know that x_1 games were played in the 2021 festival, x_2 games in 2020, up to x_t games t years ago, i.e., in the year $(2022-t)$ festival. How does the posterior PDF in part (a) change?

   Solution: The conditional PDF of X_i given Θ is Geometric(Θ), namely $f_{X_i|\Theta}(x_i|\theta) = (1-\theta)^{x_i-1}\theta$, so the posterior PDF is

$$\begin{aligned} f_{\Theta|X_1\ldots X_t}(\theta|x_1,\ldots,x_t) &\propto f_{X_1|\Theta}(x_1|\theta)\cdots f_{X_t|\Theta}(x_t|\theta) \cdot f_\Theta(\theta) \\ &= (1-\theta)^{x_1-1}\theta \cdots (1-\theta)^{x_t-1}\theta \cdot 1 \\ &= \theta^t(1-\theta)^{x_1+\cdots+x_t-t}, \end{aligned}$$

   which is the PDF of a Beta$(t+1, x_1+\cdots+x_t-t+1)$ random variable.

4. In this question you will investigate sampling using Bayesian statistics. You have a coin of an unknown probability of heads P. Your prior is that P is a Uniform$(0, 1)$ random variable.

   (a) The coin is flipped 10 times and 9 of the 10 flips are heads. What is the posterior probability that $P > 80\%$?

   Solution: The number of heads X is a Binomial$(10, P)$ random variable. The posterior PDF of P is a Beta$(10, 2)$ random variable so

$$f_{P|X}(p|9) = \frac{11!}{9!\cdot 1!}p^9(1-p) = 110p^9(1-p).$$

   The posterior probability that $P > 0.8$ is then

$$\Pr(P > 0.8|X = 9) = \int_{0.8}^{1} f_{P|X}(p|9)dp = 110\left(\frac{p^{10}}{10} - \frac{p^{11}}{11}\right)\Bigg|_{0.8}^{1} \approx 0.6779.$$

   (b) **(Optional)** A second coin, whose prior is also uniform and independent of the first coin, is flipped 10 times and all flips are heads. What is the probability that the second coin is more biased towards heads than the first one?

   Solution: Let Q and Y be the bias and number of heads of the second coin. We assume the prior of Q is Uniform$(0, 1)$ and independent of P. We are interested in the probability that $Q > P$ given that $X = 9, Y = 10$. By the total probability theorem,

$$\mathrm{P}(Q > P|X = 9, Y = 10) = \int_0^1 \Pr(Q > p|X = 9, Y = 10, P = p) \cdot f_{P|X,Y}(p|9, 10)dp.$$

Since the coins are independent and P given $X = 9$ is a $\text{Beta}(10, 2)$ random variable,

$$f_{P|X,Y}(p|9, 10) = f_{P|X}(p|9) = 110p^9(1 - p).$$

By the same reasoning, Q given $Y = 10$ is a $\text{Beta}(11, 1)$ random variable with PDF $f_{Q|Y}(q|10) = 11q^{10}$ and

$$\Pr(Q > p|X = 9, Y = 10, P = p) = \Pr(Q > p|Y = 10) = \int_p^1 11q^{10} dq = q^{11}\Big|_p^1 = 1 - p^{11}.$$

Therefore

$$\text{P}(Q > P \mid X = 9, Y = 10) = \int_0^1 (1 - p^{11}) \cdot 110p^9(1 - p) \cdot dp = \frac{16}{21} \approx 0.7619.$$