1. Suppose that $Y_1, Y_2, \cdots, Y_n$ are independent samples from a fixed Poisson distribution with an unknown mean $\lambda > 0$. Suppose that our prior distribution for λ is $\Gamma(\alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$. Compute the probability density function of the posterior distribution.

   In the following equations; let $S_n = \sum_{i=0}^{n} Y_i$

$$
\begin{aligned}
\mathbb{P}(\lambda | Y_1, \ldots, Y_n) &= \frac{\mathbb{P}(Y_1, \ldots, Y_n | \lambda)\mathbb{P}(\lambda)}{\mathbb{P}(Y_1, \ldots, Y_n)} \\
&= \frac{\frac{\lambda^{S_n} e^{-n\lambda}}{\prod_{i=1}^{n} Y_i!} \times \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)}}{\int_0^\infty \frac{x^{S_n} e^{-nx}}{\prod_{i=1}^{n} Y_i!} \times \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \,\mathrm{d}x} \\
&= \frac{\frac{\beta^\alpha}{\Gamma(\alpha)\prod_{i=1}^{n} Y_i!} \lambda^{S_n} e^{-n\lambda} \lambda^{\alpha-1} e^{-\beta\lambda}}{\frac{\beta^\alpha}{\Gamma(\alpha)\prod_{i=1}^{n} Y_i!} \int_0^\infty x^{S_n} e^{-nx} x^{\alpha-1} e^{-\beta x} \,\mathrm{d}x} \\
&= \frac{\lambda^{\alpha+S_n-1} e^{-(\beta+n)\lambda}}{\int_0^\infty x^{\alpha+S_n-1} e^{-(\beta+n)x} \,\mathrm{d}x} \\
&= \frac{\lambda^{\alpha+S_n-1} e^{-(\beta+n)\lambda}}{\frac{1}{\beta+n} \int_0^\infty \left(\frac{y}{\beta+n}\right)^{\alpha+S_n-1} e^{-y} \,\mathrm{d}y} \quad \dagger \\
&= \frac{\lambda^{\alpha+S_n-1} e^{-(\beta+n)\lambda}}{\frac{1}{(\beta+n)^{\alpha+S_n}} \int_0^\infty y^{\alpha+S_n-1} e^{-y} \,\mathrm{d}y} \\
&= \frac{(\beta+n)^{\alpha+S_n} \lambda^{\alpha+S_n-1} e^{-(\beta+n)\lambda}}{\int_0^\infty x^{\alpha+S_n-1} e^{-x} \,\mathrm{d}x} \\
&= \frac{(\beta+n)^{\alpha+S_n} \lambda^{\alpha+S_n-1} e^{-(\beta+n)\lambda}}{\Gamma(\alpha+S_n)} \\
&= \Gamma(\alpha+S_n, \beta+n)
\end{aligned}
$$

   $\dagger$: By substituting $y = (\beta+n)x$. $\frac{1}{\beta+n}\,\mathrm{d}y = \mathrm{d}x$

   So the posterior probability density function is $\Gamma(\alpha + S_n, \beta + n)$.

2. Supose Wilfred has $k > 1$ light bulbs and the probability of any individual bulb not working is p. Two strategies for testing the k bulbs are:

   (A) Test each bulb separately. This takes k tests.

   (B) Wire up all k bulbs as a series circuit. If all the bulbs come on, the testing is complete in just one test, otherwise revert to strategy A taking a total of $k + 1$ tests.

   Let X be a random variable whose value r is the number of tests required using strategy B. The probability $\mathbb{P}(X = r)$ may be expressed as:

$$
\mathbb{P}(X = r) = \begin{cases} (1-p)^k & \text{if } r = 1 \\ 1-(1-p)^k & \text{if } r = k+1 \\ 0 & \text{otherwise} \end{cases}
$$

   (a) Explain this function and justify why the constraint $k > 1$ is required mathematically.

   $X = 1$ if and only if all k bulbs work. So $\mathbb{P}(X = 1) = (1 - p)^k$ as required.

   If any bulbs don't work, then under method B we do the first test and then do a further k tests leading to $k + 1$ total tests. The probability that any bulbs don't work is $1 - \mathbb{P}(\text{all bulbs work}) = 1 - (1 - p)^k$.

No other outcomes are possible. So all other outcomes have probability zero. This leads to the function as described.

Since we are considering real objects, the constraint $k \in \mathbb{N}$ is obvious. To restrict $k > 1$ we have to consider two cases: $k = 0$ and $k = 1$.

$k = 0$ Since $k + 1 = 1$, the function claims that $\mathbb{P}(X = 1) = 0 \wedge 1$ which is a contradiction.

$k = 1$ Contraversially, I don't think the constraint $k > 1$ is required *mathematically.* It is required for method B to make any sense. If $k = 1$ then we test 1 bulb and if it's broken test it again – which wastes a test.

(b) Determine the expectation $\mathbb{E}(X)$.

$$\begin{aligned}\mathbb{E}(X) &= (1-p)^k + (k+1)(1-(1-p)^k) \\ &= k + 1 + (1-p)^k - (k+1)(1-p) \\ &= k + 1 - k(1-p)^k\end{aligned}$$

(c) Strategy B beats strategy A if $\mathbb{E}(X) < k$ and this condition is satisfied if $p < f(k)$ where $f(k)$ is some function of k. Derive the function $f(k)$.

$$\begin{aligned}k + 1 - k(1-p)^k &< k \\ 1 - k(1-p)^k &< 0 \\ 1 &< k(1-p)^k \\ \frac{1}{k} &< (1-p)^k \\ \sqrt[k]{\frac{1}{k}} &< 1 - p \\ p &< 1 - \sqrt[k]{\frac{1}{k}}\end{aligned}$$

(d) Suppose you have n light bulbs, where $n \gg k$ and $k|n.\exists m.n = m \cdot k$ and you partition the n bulbs into m groups of k. Assuming that the groups are independent and again assuming that $k < 1$, show that the expected number of tests is:

$$n\left(1 + \frac{1}{k} - (1-p)^k\right)$$

$$\begin{aligned}\mathbb{E}(X) &= m(k + 1 - k(1-p)^k) \\ &= mk\left(1 + \frac{1}{k} - (1-p)^k\right) \\ &= n\left(1 + \frac{1}{k} - (1-p)^k\right)\end{aligned}$$

In the question sheet, neither of the brackets were square. I will assume the question was asking about $(1 + \frac{1}{k} - (1-p)^k)$ rather than $(1-p)$.

If p is small then $(1 + \frac{1}{k} - (1-p)^k)$ is close to $\frac{1}{k}$.

If p is large then $(1 + \frac{1}{k} - (1-p)^k)$ is close to $1 + \frac{1}{k}$.

We can exploit this by using larger k when p is small.

The optimal value of k to choose for a given p is the solution to the equation $2 \ln k + k \ln(1-p) + \ln\left(\ln \frac{1}{1-p}\right) = 0$ which can be solved numerically.

3. You and a friend are presented with a box containing nine indistinguishable chocolates, three of which are contaminated with a deadly poison. Each of you must eat a single chocolate.

   (a) If you choose before your friend, what is the probability that you will survive?

   $$\frac{2}{3}$$

   (b) If you choose first and survive, what is the probability that your friend survives?

   $$\frac{5}{8}$$

   (c) If you choose first and die, what is the probability that your friend survives?

   $$\frac{3}{4}$$

   (d) Is it in your best interests to persuade your "friend" to choose first?

   No.

   $$\mathbb{P}(\text{survive}|\text{first}) = \frac{2}{3} \qquad \begin{aligned}\mathbb{P}(\text{survive}|\text{second}) &= \frac{1}{3} \times \frac{3}{4} + \frac{2}{3} \times \frac{5}{8} \\ &= \frac{3}{12} + \frac{5}{12} \\ &= \frac{8}{12} \\ &= \frac{2}{3}\end{aligned}$$

   (e) If you choose first, what is the probability that you survive given that your friend survives?

   $$\begin{aligned}\mathbb{P}(\text{survive}|\text{friend survives}) &= \frac{\frac{2}{3} \times \frac{5}{8}}{\frac{2}{3} \times \frac{5}{8} + \frac{1}{3} \times \frac{6}{8}} \\ &= \frac{2 \times 5}{2 \times 5 + 1 \times 6} \\ &= \frac{10}{16} \\ &= \frac{5}{8}\end{aligned}$$

4. There are n socks in a drawer, three of which are red and the rest black. Wilfred chooses his socks by selecting two at random from the drawer, and puts them on. He is three times more likely to wear socks of different colours than to wear matching red socks.

   (a) Find n.

   For legibility, I will use ${}^nC_r$ to mean $\binom{n}{r}$.

$$\begin{aligned}
\frac{\mathbb{P}(2 \text{ colours})}{\mathbb{P}(2 \text{ red})} &= \frac{\left(\frac{{}^3C_1\ {}^{n-3}C_1}{{}^nC_2}\right)}{\left(\frac{{}^3C_2\ {}^{n-3}C_0}{{}^nC_2}\right)} \\
3 &= \frac{3 \times (n-3)}{3} \\
3 &= n - 3 \\
n &= 6
\end{aligned}$$

(b) For this value of n, what is the probability that Wilfred wears matching black socks?

$$\begin{aligned}
\mathbb{P}(2 \text{ black}) &= \frac{\binom{3}{0}\binom{3}{2}}{\binom{6}{2}} \\
&= \frac{3}{15} \\
&= \frac{1}{5}
\end{aligned}$$

5. Let $X_1, X_2, \ldots$ be a sequence of observations. Define the sample mean $\overline{X}_n$, of the first n observations by $\overline{X}_n = \sum_{i=1}^{n} \frac{X_i}{n}$. Show that for $n > 1$

$$\overline{X}_n = \overline{X}_{n-1} + \frac{X_n - \overline{X}_{n-1}}{n}$$

$$\begin{aligned}
\overline{X}_n &= \frac{\sum_{i=1}^{n} X_i}{n} \\
&= \frac{(n-1)\overline{X}_{n-1} + X_n}{n} \\
&= \frac{n\overline{X}_{n-1} + X_n - \overline{X}_{n-1}}{n} \\
&= \overline{X}_{n-1} + \frac{X_n - \overline{X}_{n-1}}{n}
\end{aligned}$$

6. A canon fixed at the origin in the $x - y$ plane fires discrete particles. The i^{th}] particle moves within the plane along a straight line which makes a random angle $\Theta = \theta_i$ with the $x - axis$ and collides with a screen at (x_0, y_i) where the parameter $x_0 > 0$ defines the position of the screen on the x-axis. The values of θ_i are uniformly distributed in the range $-\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}$.

(a) Find y_i in terms of θ_i and x_0.

$$y_i = x_0 \tan \theta_i$$

(b) Find the probability density function, $g(\theta)$ and cumulative density function $G(\theta) = \mathbb{P}(\Theta \leq \theta)$.

$$g(\theta) = \frac{1}{\pi}$$

$$G(\theta) = \frac{\theta + \frac{\pi}{2}}{\pi}$$

(c) Letting Y be a continuous random variable denoting the y-coordinate of the collision points. Find the cumulative distribution function $F(y) = \mathbb{P}(Y \leq y)$.

$$F(y) = \frac{\arctan\left(\frac{y}{x_0}\right) + \frac{\pi}{2}}{\pi}$$

(d) Hence show that the probability distribution function $f(y)$ is given by:

$$f(y) = \frac{x_0}{\pi(x_0^2 + y^2)}$$

$$\begin{aligned} f(y) &= \frac{\mathrm{d}F}{\mathrm{d}y} \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\pi}\arctan\left(\frac{y}{x_0}\right) + \frac{1}{2}\right) \\ &= \frac{x_0}{\pi(x_0^2 + y^2)} \end{aligned}$$

(e) Show that the standard deviation of the y-coordinate of the collision position has no mathematically defined value.

$$\begin{aligned} Var(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\ &= \int_{-\infty}^{\infty} \frac{x_0 y^2}{\pi(x_0^2 + y^2)}\,\mathrm{d}x - \left(\int_{-\infty}^{\infty} \frac{x_0 y}{\pi(x_0^2 + y^2)}\,\mathrm{d}y\right) \\ &= \left[x_0 y - x_0^2 \arctan\left(\frac{y}{x_0}\right)\right]_{-\infty}^{\infty} - \left[\frac{1}{2}x_0 \ln\left(y^2 + x_0^2\right)\right]_{-\infty}^{\infty} \\ &= (\infty - x_0^2\pi + \infty) - (\infty - \infty) \\ &= \infty - \infty \\ &\text{Which is undefined} \end{aligned}$$

(f) θ is now uniformly distributed between $-\frac{\pi}{6}$ and $+\frac{\pi}{3}$. Find $\mathbb{E}(Y)$.

$$G(y) = \frac{2}{\pi}\arctan\left(\frac{y}{x_0}\right) + \frac{1}{3}$$

$$\begin{aligned} g(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\left(\frac{2}{\pi}\arctan\left(\frac{y}{x_0}\right) + \frac{1}{3}\right) \\ &= \frac{2x_0}{\pi(x_0^2 + y^2)} \end{aligned}$$

$$
\begin{aligned}
\mathbb{E}(y) &= \int_{-\frac{x_0}{\sqrt{3}}}^{\sqrt{3}x_0} \frac{2x_0 y}{\pi(x_0^2+y^2)}\,\mathrm{d}y \\
&= \left[\frac{x_0}{\pi}\ln\left(x_0^2+y^2\right)\right]_{-\frac{x_0}{\sqrt{3}}}^{\sqrt{3}x_0} \\
&= \frac{x_0}{\pi}\left(\ln\left(4x_0^2\right) - x_0\ln\left(\frac{4}{3}x_0^2\right)\right) \\
&= \frac{x_0}{\pi}\ln(3)
\end{aligned}
$$

(g) The distance to the screen now halves instantaneously every time a particle is fired. Find $\mathbb{E}\left(\sum_{i=1}^{\infty} Y_i\right)$

$$
\begin{aligned}
\mathbb{E}\left(\sum_{i=1}^{\infty} Y_i\right) &= \sum_{i=1}^{\infty}\mathbb{E}(Y_i) \\
&= \frac{x_0\ln 3}{\pi}\sum_{i=1}^{\infty}\frac{1}{2}^{i} \\
&= \frac{x_0\ln 3}{\pi}
\end{aligned}
$$