Exercise 1 (chapter 1.4): Conditionally on $\Theta = \theta$, $\{Y_i\}_{i=1}^n$ are independent and identically distributed random variables that follow an exponential distribution with rate θ. The density of the exponential distribution is

$$p(y \mid \theta) = \theta \exp\{-\theta y\}\,, \qquad y > 0\,.$$

Let the prior on Θ be a Gamma distribution with shape $\alpha = 1$ and rate $\beta = 1$. There are two datasets:

1. $n = 5$ and $\bar{y} = n^{-1}\sum_{i=1}^n y_i = 0.25$
2. $n = 100$ and $\bar{y} = 0.25$

For both datasets, plot the prior, likelihood, the product of prior and likelihood as well as the posterior density (which happens to be a Gamma density).

Solution: The density of the Gamma prior on Θ is

$$p(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}\exp\{-\beta\theta\} \qquad \theta, \alpha, \beta > 0\,.$$

The likelihood of Θ is

$$p(y \mid \theta) = \prod_{i=1}^n \theta \exp\{-\theta y_i\} = \theta^n \exp\{\theta n \bar{y}\}\,.$$

Combining the prior density and likelihood, the posterior density of Θ is proportional to

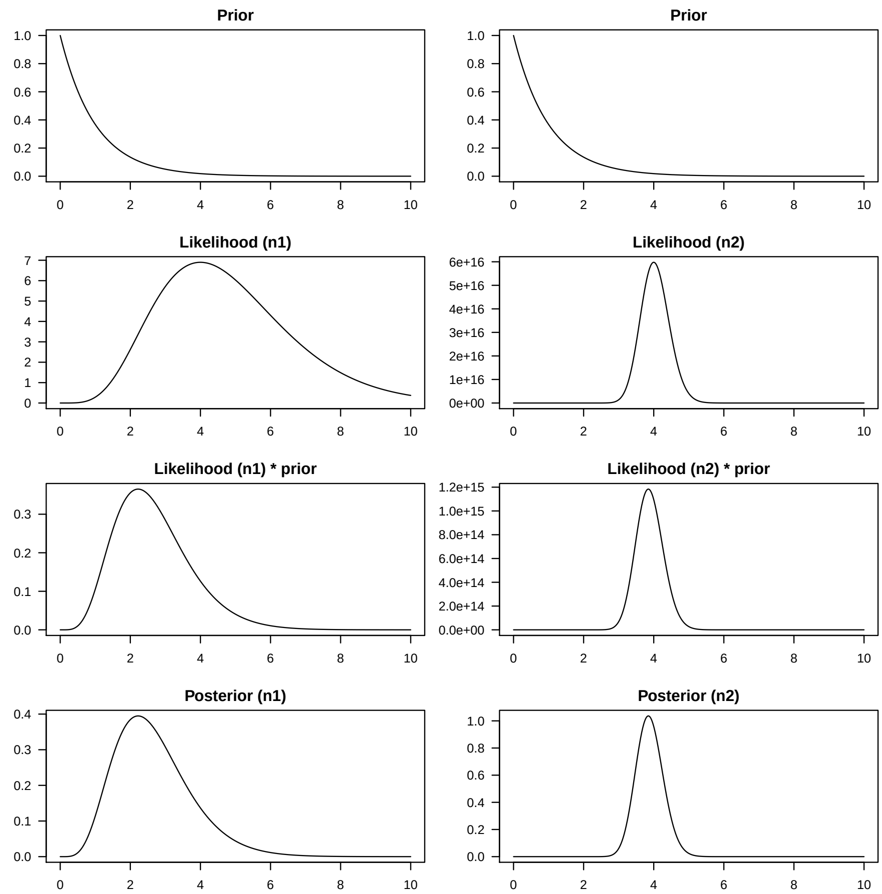
$$p(\theta \mid y) \propto p(y \mid \theta)p(\theta) = \theta^n \exp\{\theta n\bar{y}\}\theta^{\alpha-1}\exp\{-\beta\theta\} = \theta^{\alpha+n-1}\exp\{-\theta(\beta + n\bar{y})\}\,,$$

which represents the kernel of a $\text{Gamma}(\theta \mid \alpha + n, \beta + n\bar{y})$ distribution.

Exercise 2 (chapter 1.4): For the statistical model from Exercise 1, find a closed form formula for the predictive density

$$p(y^* \mid y) = \int_\Theta p(y^*, \theta \mid y)\,\mathrm{d}\theta = \int_\Theta p(y^* \mid \theta)p(\theta \mid y)\,\mathrm{d}\theta$$

of a new observation y^*. Evaluate and plot the predictive density for the first dataset from Exercise 1 by setting up a grid for the y^* values.



Solution: The predictive density is

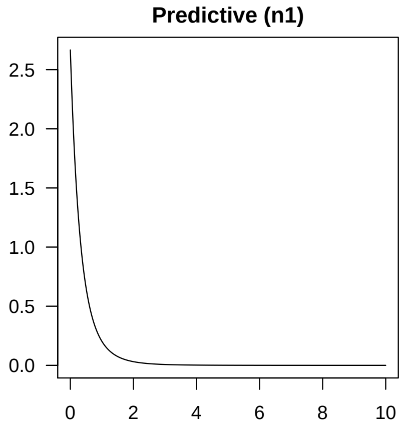
$$
\begin{aligned}
p(y^* \,|\, y) &= \int_0^\infty \theta \exp\{-\theta y^*\} \frac{(\beta + n\bar{y})^{\alpha+n}}{\Gamma(\alpha+n)} \theta^{\alpha+n-1} \exp\{-\theta(\beta + n\bar{y})\} \, \mathrm{d}\theta \\
&= \frac{(\beta + n\bar{y})^{\alpha+n}}{\Gamma(\alpha+n)} \int_0^\infty \theta^{\alpha+n} \exp\{-\theta(\beta + y^* + n\bar{y})\} \, \mathrm{d}\theta \\
&= \frac{(\beta + n\bar{y})^{\alpha+n}}{\Gamma(\alpha+n)} \frac{\Gamma(\alpha+n+1)}{(\beta + y^* + n\bar{y})^{\alpha+n+1}} \\
&= \frac{(\alpha+n)(\beta + n\bar{y})^{\alpha+n}}{(\beta + y^* + n\bar{y})^{\alpha+n+1}} \,,
\end{aligned}
$$

where the integral in the second line is the inverse of the normalizing constant of a Gamma distribution and the last line used the following property of the Gamma function: $\Gamma(x+1) = x\Gamma(x)$.

Exercise 3 (chapter 2.7): The joint conditional distribution of Y^* and Θ factorizes as

$$p(y^*, \theta \,|\, y) = p(y^* \,|\, \theta) p(\theta \,|\, y) \,,$$

because the random variables Y and Y^* are conditionally independent given $\Theta = \theta$. Derive this results from the multiplication rule for conditional distributions.



Solution: Using the multiplication rule and conditional independence of Y and Y^* given $\Theta = \theta$ gives

$$p(y^*, \theta \mid y) = p(y^* \mid y, \theta)p(\theta \mid y) = p(y^* \mid \theta)p(\theta \mid y)\,.$$

Exercise 4 (chapter 2.10): Let the random variable X follow a Gamma distribution with shape $\alpha > 0$ and rate $\beta > 0$. There is only information about $Y = g(X) = X^{-1}$. The distribution of Y is the Inverse-Gamma distribution with parameters α and β.

1. Find the density of Y using a change-of-variables:
$$f_Y(y) = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = f_X(h(y))|h'(y)| \text{ under the bijection } y = g(x) \Leftrightarrow x = h(y)$$
2. Find a formula for the mode (i.e. the maximum point) of the density of Y
3. Find the expectation $\mathbb{E}[Y]$ assuming $\alpha > 1$ using $\mathbb{E}\big[X^{-1}\big]$

Solution: The density of Y is

$$f_Y(y) = f_X(x)\left|\frac{\mathrm{d}x}{\mathrm{d}y}\right| = \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{y}\right)^{\alpha-1}\exp\left\{-\frac{\beta}{y}\right\}\frac{1}{y^2} = \frac{\beta^\alpha}{\Gamma(\alpha)}y^{-\alpha-1}\exp\{-\beta/y\} \qquad y > 0\,.$$

The mode of $f_Y(y)$ is

$$\frac{\mathrm{d}\log f_Y(y)}{\mathrm{d}y} = -\frac{\alpha+1}{y} + \frac{\beta}{y^2} \overset{!}{=} 0 \Rightarrow y = \frac{\beta}{\alpha+1}\,.$$

The second derivative is

$$\frac{\mathrm{d}^2\log f_Y(y)}{\mathrm{d}y^2} = \frac{\alpha+1}{y^2} - \frac{2\beta}{y^3}\,,$$

which is negative at $y = \beta/(\alpha+1)$ showing that it is a maximum point. The expectation of Y is

$$\begin{aligned}
\mathbb{E}[Y] = \mathbb{E}\big[X^{-1}\big] &= \int_0^\infty \frac{1}{x}\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-\beta x\}\, \mathrm{d}x \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{(\alpha-1)-1} \exp\{-\beta x\}\, \mathrm{d}x \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-1)}{\beta^{\alpha-1}} = \frac{\beta}{\alpha-1} \qquad \alpha > 1\,.
\end{aligned}$$

Exercise 5 (chapter 2.10): Let the random variables $\{X_i\}_{i=1}^3$ follow independently Gamma distributions with shape $\alpha_1, \alpha_2, \alpha_3 > 0$ and rate $\beta_1 = \beta_2 = \beta_3 = 1$. Using a multivariate change-of-variables

$$Y_1 = \frac{X_1}{X_1 + X_2 + X_3} \qquad Y_2 = \frac{X_2}{X_1 + X_2 + X_3} \qquad S = X_1 + X_2 + X_3\,,$$

find the joint density of Y_1, Y_2 and S. Find also the joint density of Y_1 and Y_2 by integrating out S (which happens to be a Dirichlet distribution).

Solution: The change-of-variables gives

$$X_1 = Y_1 S \qquad X_2 = Y_2 S \qquad X_3 = S(1 - Y_1 - Y_2)\,.$$

The determinant of the Jacobian matrix is required for the change-of-variable:

$$\left|\frac{\partial x_1, x_2, x_3}{\partial y_1, y_2, s}\right| = \det\begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \frac{\partial x_1}{\partial s} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \frac{\partial x_2}{\partial s} \\ \frac{\partial x_3}{\partial y_1} & \frac{\partial x_3}{\partial y_2} & \frac{\partial x_3}{\partial s} \end{bmatrix} = \det\begin{bmatrix} s & 0 & y_1 \\ 0 & s & y_2 \\ -s & -s & 1 - y_1 - y_2 \end{bmatrix} = s^2\,.$$

The joint density of Y_1, Y_2 and S is then

$$\begin{aligned}
f_{Y_1,Y_2,S}(y_1, y_2, s) &= f_{X_1,X_2,X_3}(x_1, x_2, x_3)\left|\frac{\partial x_1, x_2, x_3}{\partial y_1, y_2, s}\right| \\
&= \frac{1}{\Gamma(\alpha_1)}(y_1 s)^{\alpha_1 - 1} e^{-y_1 s} \times \\
&\quad \frac{1}{\Gamma(\alpha_2)}(y_2 s)^{\alpha_2 - 1} e^{-y_2 s} \times \\
&\quad \frac{1}{\Gamma(\alpha_3)}(s[1 - y_1 - y_2])^{\alpha_3 - 1} e^{-(1-y_1-y_2)s} \times \\
&\quad s^2 \\
&= \frac{y_1^{\alpha_1-1} y_2^{\alpha_2-1} (1 - y_1 - y_2)^{\alpha_3-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} s^{\alpha_1+\alpha_2+\alpha_3-1} e^{-s}\,.
\end{aligned}$$

The marginal density of Y_1 and Y_2 is

$$
\begin{aligned}
f_{Y_1,Y_2}(y_1,y_2) &= \int_0^\infty f_{Y_1,Y_2,S}(y_1,y_2,s)\,\mathrm{d}s \\
&= \frac{y_1^{\alpha_1-1}y_2^{\alpha_2-1}(1-y_1-y_2)^{\alpha_3-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} \int_0^\infty s^{\alpha_1+\alpha_2+\alpha_3-1}e^{-s}\,\mathrm{d}s \\
&= \frac{\Gamma(\alpha_1+\alpha_2+\alpha_3)}{\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3)} y_1^{\alpha_1-1}y_2^{\alpha_2-1}(1-y_1-y_2)^{\alpha_3-1} \qquad y1, y2 > 0 \text{ and } 0 < y_1+y_2 < 1\,,
\end{aligned}
$$

where the integral in the second line is the inverse of the normalizing constant of a Gamma distribution. The marginal density is that of a Dirichlet distribution with parameters α_1, α_2 and α_3.

Exercise 6 (chapter 3.2): Let the random variable X follow a Pareto distribution with shape $\alpha > 0$ and scale $x_m > 0$. The density of the Pareto distribution is

$$
f_X(x) = \frac{\alpha\, x_m^\alpha}{x^{\alpha+1}}\,, \qquad x \geq x_m\,.
$$

Derive the inverse transformation method to simulate from the Pareto distribution (there is no function in the standard packages of R).

Solution: The cumulative distribution function of X is

$$
F_X(x) = 1 - \left(\frac{x_m}{x}\right)^\alpha \qquad x \geq x_m\,.
$$

The quantile function $F_X^{-1}(u)$ is now straightforward to derive

$$
F_X^{-1}(u) = \frac{x_m}{(1-u)^{1/\alpha}} \qquad 0 < u < 1\,.
$$

The inverse transform method is then

$$
\begin{aligned}
&\text{Simulate } U \sim \text{Unif}(0,1) \\
&\quad\text{Set } X = \frac{x_m}{U^{1/\alpha}}\,.
\end{aligned}
$$

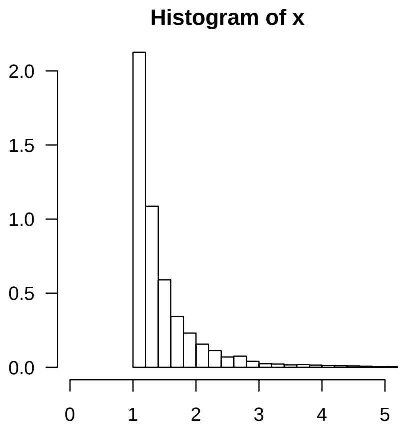
A sample of 10000 random numbers from the Pareto distribution can be generated in R.

Exercise 7 (chapter 3.4): Let $f_X(x)$ be the density of a continuously distributed random variable X. The cumulative distribution $F_X(x)$ and quantile function $F_X^{-1}(u)$ with $u \in (0,1)$ are known. Derive the inverse transformation method when the distribution of X is truncated to the interval $I = (a,b)$ with $a < b$. The density of the truncated distribution is proportional to the unnormalized density

$$
g_X^*(x) \propto f_X(x) 1_{(a,b)}(x)\,.
$$

Start by determining the normalizing constant k such that $g_X(x) = g_X^*(x)/k$ is a density and then derive

```
alpha <- 3
xm <- 1
x <- xm/runif(10000)^(1/alpha)
par(mar = c(3, 3, 2, 2), las = 1)
hist(x, breaks = "Scott", probability = TRUE, xlim = c(0, 5), xlab = "", ylab = "")
```



the cumulative distribution $G_X(x)$ and quantile function $G_X^{-1}(u)$ of the truncated distribution.

Solution: The normalizing constant of $g_X(x)$ is

$$1 = \int_{-\infty}^{\infty} g_X(x)\,\mathrm{d}x = \frac{1}{k}\int_{-\infty}^{\infty} g_X^*(x)\,\mathrm{d}x = \frac{1}{k}\int_a^b f_X(x)\,\mathrm{d}x = \frac{F_X(b) - F_X(a)}{k} \Rightarrow k = F_X(b) - F_X(a)\,.$$

For $a < x < b$, the cumulative distribution function is

$$G_X(x) = \int_a^x g_X(t)\,\mathrm{d}t = \frac{1}{F_X(b) - F_X(a)}\int_a^x f_X(t)\,\mathrm{d}t = \frac{F_X(x) - F_X(a)}{F_X(b) - F_X(a)}\,.$$

The quantile function is now straightforward to derive

$$G_X^{-1}(u) = F_X^{-1}\{F_X(a) + u[F_X(b) - F_X(a)]\} \qquad 0 < u < 1\,.$$

The inverse transform method is then

Simulate $U \sim \mathrm{Unif}(0,1)$

Set $X = F_X^{-1}\{F_X(a) + U[F_X(b) - F_X(a)]\}\,.$