

Assignment #2 Solutions

1. A brand of lightbulbs is advertised to have a mean life of 1000 hours.

Solution:

- (a) Find t so that $P(T < t) = 0.95$ when $T \sim \text{Exponential}(1/1000)$.

If the mean is 1000 hours, this implies that $\lambda = 1/1000 = 0.001$. Since the exponential density is decreasing, the shortest interval will be of the form $(0, t)$, and t will be the 0.95 quantile of the distribution.

The exponential distribution has right-tail area $e^{-\lambda t}$, so $0.05 = P(T > t) = e^{-t/1000}$ which is solved when $t = -1000 \log(0.05) \doteq 2995.7$.

More simply using R,

```
> qexp(0.95, 1/1000)
```

```
[1] 2995.732
```

- (b) Find t so that $P(T > t) = 0.9$ when $t \sim \text{Gamma}(4, \lambda)$ and $E(T) = 1000$.

Since $\mu = \alpha/\lambda = 4/\lambda = 1000$, $\lambda = 4/1000 = 1/250 = 0.004$. Let F be the cumulative distribution function for the $\text{Gamma}(4, 0.004)$ distribution. We can find t as follows:

$$\begin{aligned} P(T > t) &= 0.9 \\ 1 - P(T > t) &= 0.1 \\ F(t) &= 0.1 \\ t &= F^{-1}(0.1) \\ &\doteq 436.2 \end{aligned}$$

The last equality uses the following R calculation.

```
> qgamma(0.1, 4, 0.004)
```

```
[1] 436.1924
```

- (c) Find t so that $P(T > 500 + t | T > 500) = 0.9$ when $t \sim \text{Gamma}(4, \lambda)$ and $E(T) = 1000$.

$$\begin{aligned} P(T > 500 + t | T > 500) &= 0.9 \\ \frac{P(T > 500 + t)}{P(T > 500)} &= 0.9 \\ \frac{1 - F(500 + t)}{1 - F(500)} &= 0.9 \\ 1 - F(500 + t) &= 0.9(1 - F(500)) \\ F(500 + t) &= 1 - 0.9(1 - F(500)) \\ t &= F^{-1}(1 - 0.9(1 - F(500))) - 500 \\ &\doteq 108.7 \end{aligned}$$

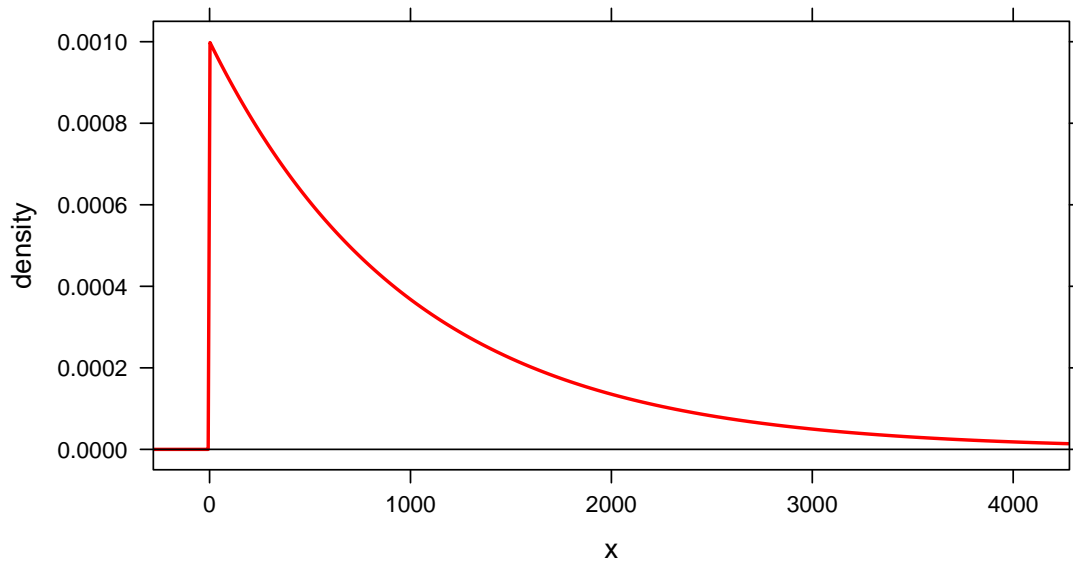
The last equality uses the following R calculation.

```
> qgamma(1 - 0.9 * (1 - pgamma(500, 4, 0.004)), 4, 0.004) - 500
```

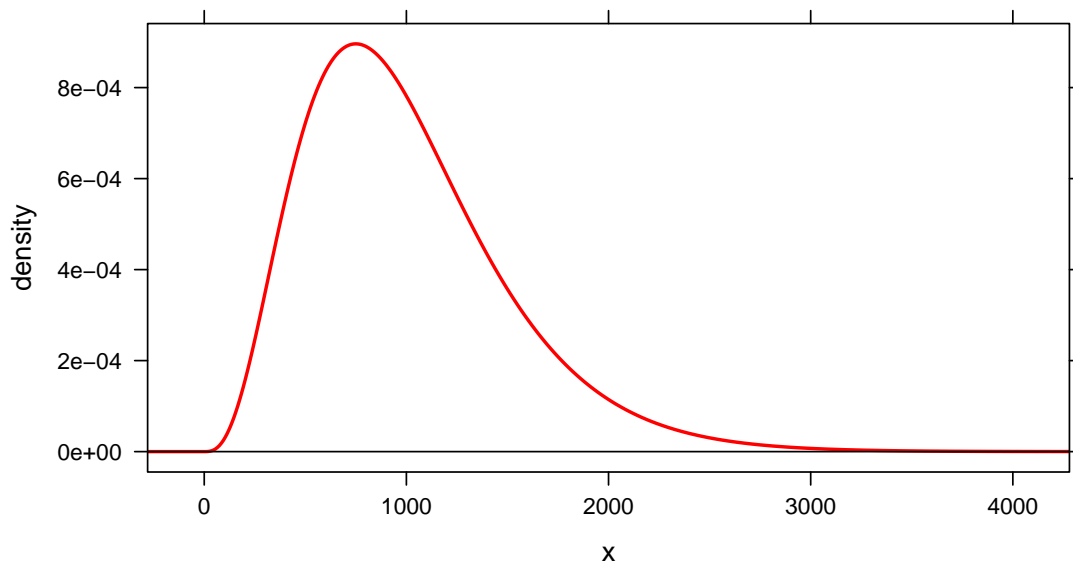
```
[1] 108.7085
```

- (d) Graph both the exponential and gamma densities from this problem. The following examples assumes that the files prob.R has been sourced.

```
> fig1 = gexp(endpoints = c(0, 4000), 1/1000)
> print(fig1)
```



```
> fig2 = ggamma(endpoints = c(0, 4000), 4, 4/1000)
> print(fig2)
```



2. Suppose that a random variable $X \sim \text{Beta}(\alpha, \beta)$ and assume that $\alpha, \beta > 1$.

Solution:

(a) Find the mean (expected value) and mode (point that maximizes $f(x)$) in terms of α and β .

$$\begin{aligned} E(X) &= \int_0^1 x \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^{(\alpha+1)-1} (1-x)^{\beta-1} dx \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(\alpha + 1)\Gamma(\beta)}{\Gamma(\alpha + \beta + 1)} \\ &= \frac{\alpha}{\alpha + \beta} \end{aligned}$$

We find the mode by first setting the derivative of the natural logarithm of the density to zero and solving.

$$\begin{aligned} f(x) &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \\ \log f(x) &= \log \left(\frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \right) + (\alpha - 1) \log(x) + (\beta - 1) \log(1 - x) \\ \frac{d \log f(x)}{dx} &= \frac{\alpha - 1}{x} - \frac{\beta - 1}{1 - x} = 0 \\ \frac{\alpha - 1}{x} &= \frac{\beta - 1}{1 - x} \\ (\alpha - 1)(1 - x) &= (\beta - 1)x \\ x &= \frac{\alpha - 1}{\alpha + \beta - 2} \end{aligned}$$

Note that this solution requires is valid since $\alpha > 1, \beta > 1$ so that $\alpha + \beta > 2$.

(b) Assume that $\alpha = 2$ and $\beta = 3$. Assess the accuracy of a predicted value x_0 for X by calculating the mean squared error (MSE),

$$E((X - x_0)^2) = \int_0^1 (x - x_0)^2 f(x) dx .$$

Think of the MSE as the expected penalty for a wrong guess of X where the guess is x_0 and the penalty is $(X - x_0)^2$. The guess x_0 is better than the guess x_1 on average (by the squared error criterion) if its MSE is smaller. We evaluate the MSE for an arbitrary guess x_0 as follows, letting

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)} = \int_0^1 x^{a-1} (1-x)^{b-1} dx$$

be the Beta function. Recall as well that $\Gamma(a + 1) = a\Gamma(a)$.

$$\begin{aligned} E((X - x_0)^2) &= \int_0^1 (x - x_0)^2 f(x) dx \\ &= \int_0^1 (x^2 - 2x_0x + x_0^2) f(x) dx \\ &= \int_0^1 (x^2 - 2x_0x + x_0^2) f(x) dx \\ &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} (B(\alpha + 2, \beta) - 2x_0B(\alpha + 1, \beta) + x_0^2B(\alpha, \beta)) \\ &= \frac{(\alpha + 1)\alpha}{(\alpha + \beta + 1)(\alpha + \beta)} - 2x_0\frac{\alpha}{\alpha + \beta} + x_0^2 \end{aligned}$$

Plugging in $\alpha = 2, \beta = 3$ results in

$$E((X - x_0)^2) = \frac{1}{5} - \frac{4x_0}{5} + x_0^2$$

For the mean, $x_0 = \alpha/(\alpha + \beta) = 2/5$ and the MSE is

$$\frac{1}{5} - \frac{8}{25} + \frac{4}{25} = \frac{1}{25} = 0.04 .$$

For the mode, $x_0 = (\alpha - 1)/(\alpha + \beta - 2) = 1/3$ and the MSE is

$$\frac{1}{5} - \frac{4}{15} + \frac{1}{9} = \frac{2}{45} \doteq 0.0444 .$$

The mean is better. In fact, it is true in general that if the variance of a probability distribution is finite, the mean is the statistic that minimizes the MSE.

3. Do Exercise 5.3.1.

Solution:

In a sequence of five Bernoulli trials, the data is of the form $S = \{s = (x_1, \dots, x_5) : x_i \in \{0, 1\}\}$ where $x_i = 1$ if the i th coin toss is a head.

For a single response, the data is $s = x_1 \in \{0, 1\}$.

The likelihood model for a single response is

$$f_\theta(s) = \theta^{x_1}(1 - \theta)^{1-x_1} .$$

The likelihood model for the sample is

$$f_\theta(s) = \prod_{i=1}^5 \theta^{x_i}(1 - \theta)^{1-x_i} = \theta^{\sum_{i=1}^5 x_i}(1 - \theta)^{5-\sum_{i=1}^5 x_i} .$$

The parameter space is $\Omega = \{1/3, 1/2, 2/3\}$ for both the single response and the sample.

4. Do Exercise 5.3.3.

Solution: The data is sampled from one of two normal populations, either I where $N(10, 2)$ or II where $N(8, 3)$. We describe the statistical model.

The form of the data is $s = (x_1, \dots, x_n) \in \mathbb{R}^n$.

The probability density is

$$\begin{aligned} f_{\theta}(s) &= \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i-\mu}{\sigma}\right)^2} \\ &= (2\pi)^{-n/2} \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i-\mu)^2} \\ &= (2\pi)^{-n/2} \sigma^{-n} e^{-\frac{n}{2\sigma^2} (\hat{\sigma}^2 + (\bar{x}-\mu)^2)} \end{aligned}$$

where

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

The parameter is $\theta = (\mu, \sigma^2)$.

The parameter space is $\Omega = \{(10, 2), (8, 3)\}$.

It is also possible to index the models by $\theta \in \{I, II\}$, or $\theta \in \{10, 8\}$, or $\theta \in \{2, 3\}$, or $\theta \in \{\text{jellyfish}, \text{donut}\}$. A single index must specify a single probability measure completely. In this problem with only two models to choose from, any two distinct labels suffices. Regardless of what labels are used, it is necessary to specify which specific model the label refers to. The choice I made is particularly clear and succinct, but not unique in its validity.

5. Do Exercise 5.3.8.

Solution: Reparameterize Example 5.3.4 so that the parameter is $\theta = (\mu, \psi) = (\mu, \log(\sigma))$.

Note that if $\psi = \log(\sigma)$, then $\sigma = e^{\psi}$.

The form of the data is $s = (x_1, \dots, x_n) \in \mathbb{R}^n$.

The probability density is

$$\begin{aligned} f_{\theta}(s) &= (2\pi)^{-n/2} \sigma^{-n} e^{-\frac{n}{2\sigma^2} (\hat{\sigma}^2 + (\bar{x}-\mu)^2)} \\ &= (2\pi)^{-n/2} e^{-n\psi} e^{-\frac{n}{2e^{2\psi}} (\hat{\sigma}^2 + (\bar{x}-\mu)^2)} \end{aligned}$$

where

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad \text{and} \quad \hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

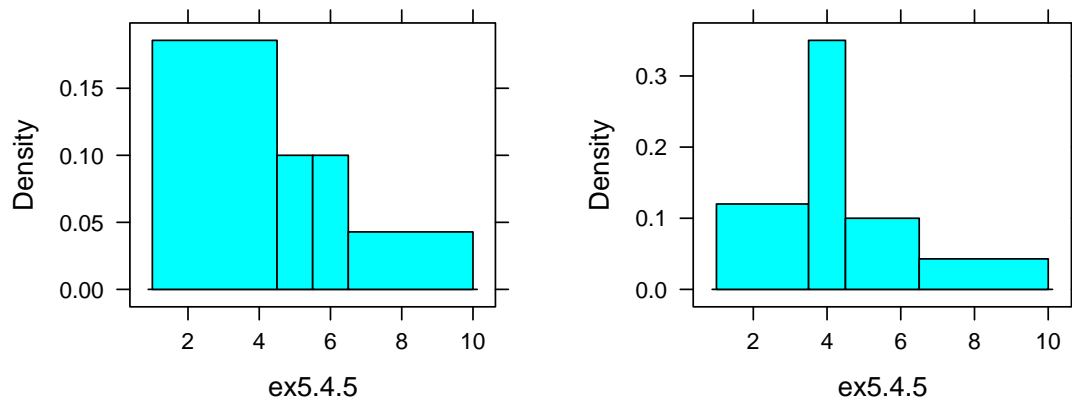
The parameter space is $\Omega = \mathbb{R} \times \mathbb{R}$.

This type of parameterization may be useful, for example, when using a canned optimization routine to seek parameter values that maximize the likelihood. A canned routine may not be able to handle the restriction $\sigma > 0$, but the reparameterized model can be optimized over \mathbb{R}^2 , and then the 1-1 correspondence can be used to recover the estimates in the original parameterization.

6. Do Exercise 5.4.5.

```
> ex5.4.5 = c(3.9, 7.2, 6.9, 4.5, 5.8, 3.7, 4.4, 4.5, 5.6, 2.5,
+ 4.8, 8.5, 4.3, 1.2, 2.3, 3.1, 3.4, 4.8, 1.8, 3.7)

> figA = histogram(~ex5.4.5, breaks = c(1, 4.5, 5.5, 6.5, 10),
+ type = "density")
> figB = histogram(~ex5.4.5, breaks = c(1, 3.5, 4.5, 6.5, 10),
+ type = "density")
> easyLayout(list(figA, figB), 1, 2)
```



Decisions on how to select breaks for histograms can greatly affect the interpretation of the distribution visually. Another option is to use `densityplot()` from the lattice package.

```
> print(densityplot(~ex5.4.5))
```

