5.1 Likelihood: Practical Questions - solutions

Question 1

The number of customers in a supermarket during lunch hour, X, is a non-negative, discrete random variable. The **Poisson** distribution is often used to model such a variable, and is defined by a single rate parameter λ. The store manager wishes to estimate the number of customers to expect during lunch hour. On one particular day, they count 26 customers between 12pm and 1pm.

(a) Write down the probability density function of the Poisson distribution.

$$P(X = x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

(b) Write down the likelihood function based on your answer in part (a), given the observed data $x = 26$.

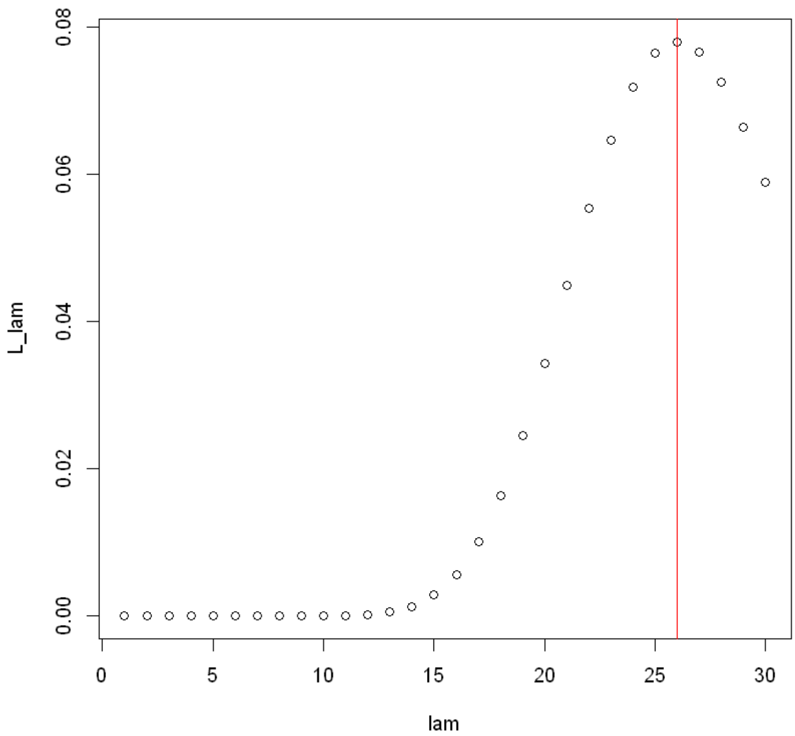
$$L(\lambda|x = 26) = \frac{\lambda^{26} e^{-\lambda}}{26!}$$

(c) Following the example from the lecture and the prompts below, calculate the likelihood for a range of values for λ. Draw a plot of the results.

(d) Identify the value of λ for which the probability of observing 26 customers is highest, i.e. for which the likelihood is highest. Mark it on your plot.

In [2]:

```
# Define range of possible values for lambda
lam <- seq(1:30)

# Calculate the Poisson likelihood for each value of lambda
L_lam <- lam^26*exp(-lam)/factorial(26)

# Plot the likelihood
plot(x = lam, y = L_lam)

# Find the MLE
mle <- lam[which.max(L_lam)]

# Indicate MLE on plot
abline(v = mle, col = "red")
```



(e) Derive the maximum likelihood estimator for λ algebraically, following the steps in section 5.6 of the lecture notes. Confirm that your estimate from part (d) agrees with the algebraic solution.

We have the likelihood for lambda

$$L(\lambda|x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

and so the log-likelihood is

$$l(\lambda|x) = x\log\lambda - \lambda - \log x!$$

Differentiating with respect to lambda

$$\frac{dl(\lambda|x)}{d\lambda} = \frac{x}{\lambda} - 1 = 0$$
$$\implies \hat{\lambda} = x$$

Finally we check the second derivative,

$$\frac{dl^2(\lambda|x)}{d\lambda^2} = -\frac{x}{\lambda^2}$$

This expression will be negative for any value of λ we substitute, so we have found our MLE.

Question 2

Recall the example from the lecture about waiting time at a GP, which we modelled using an exponential distribution. The time taken for an event to occur can also be modelled using a **Gamma** distribution.

This distribution two parameters in general: a shape parameter α and a rate parameter β. We will consider the case where α is *known* and equal to 2. We want to estimate the unknown parameter, β.

The PDF for a Gamma distribution with $\alpha = 2$ is

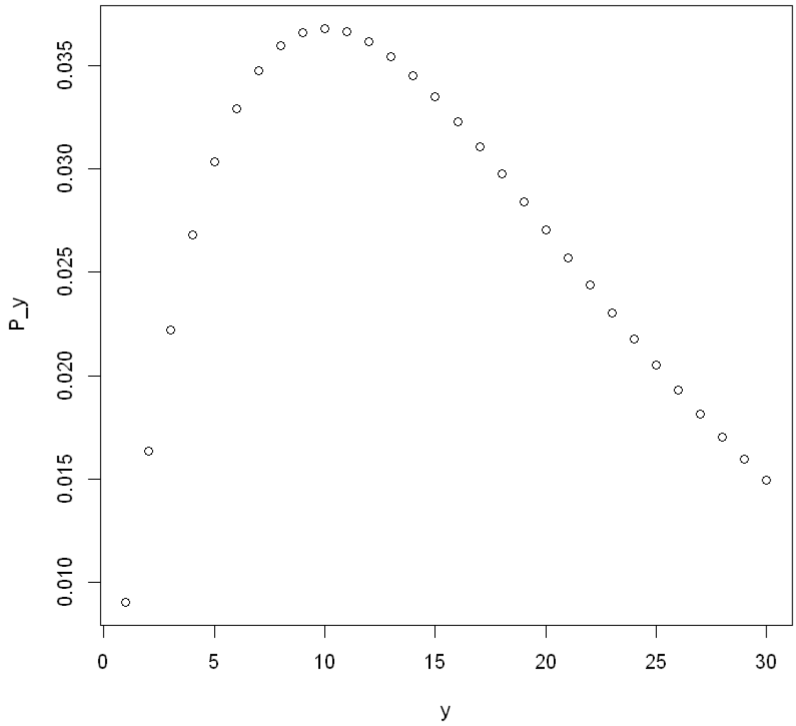
$$P(Y = y) = \frac{1}{\Gamma(2)}\beta^2 y e^{-\beta y}$$

You might notice that the PDF is similar to that of the exponential distribution. In fact, the exponential is a special case of the Gamma, when $\alpha = 1$.

(a) Plot the PDF for a Gamma distribution with $\alpha = 2$ and $\beta = 0.1$.

In [3]:

```
y <- 1:30
P_y <- dgamma(y, shape = 2, rate = 0.1)
plot(y, P_y)
```



(b) Write down the likelihood for β, given some observed data $Y = y$ and assuming $\alpha = 2$ is known.

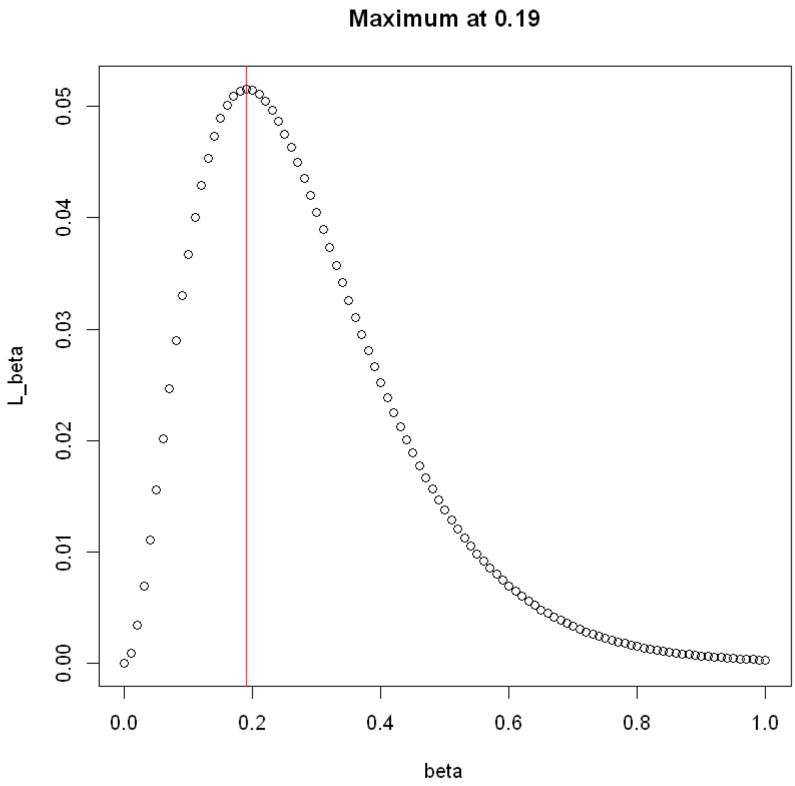
Solution: The likelihood is

$$L(\beta) = \frac{1}{\Gamma(2)}\beta^2 y e^{-\beta y}$$

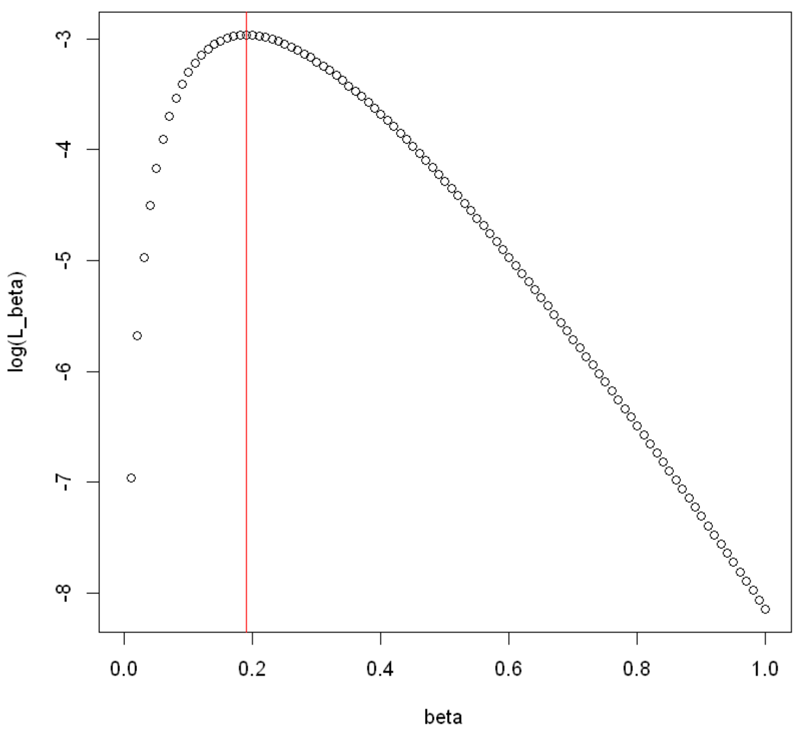
(c) Given $\alpha = 2$ and observed data $y = 10.5$, plot the likelihood and log-likelihood for some possible values of β. How does the likelihood differ from the PDF? Estimate the MLE from your plots.

In [6]:

```
# Range for beta
beta <- seq(0,1,0.01)

# Gamma likelihood for each value of beta (have a look at help(dgamma))
L_beta <- dgamma(10.5, shape = 2, rate = beta)

# Find the MLE
mle <- beta[which.max(L_beta)]

# Plot the likelihood
plot(beta, L_beta, main = paste0("Maximum at ",mle))
# Indicate MLE on plot
abline(v = mle, col = "red")

# Plot the log-likelihood
plot(beta, log(L_beta))
abline(v = mle, col = "red")
```



Maximum at 0.19
L_beta
0.00
0.01
0.02
0.03
0.04
0.05
0.0
0.2
0.4
0.6
0.8
1.0
beta


log(L_beta)
-3
-4
-5
-6
-7
-8
0.0
0.2
0.4
0.6
0.8
1.0
beta

(d) Derive the general form of the MLE for β algebraically.

Solution: The log-likelihood is

$$l(\beta|y) = \log\left(\frac{1}{\Gamma(2)}\right) + 2\log\beta + \log y - \beta y$$

Deriving with respect to β and setting to zero,

$$\frac{dl(\beta|y)}{d\beta} = \frac{2}{\beta} - y = 0$$
$$\frac{2}{\beta} = y$$
$$\hat{\beta} = \frac{2}{y}$$

Finding the second derivative,

$$\frac{dl^2(\beta|y)}{d\beta^2} = -\frac{2}{\beta^2}$$

which is negative for any value of β therefore we have our MLE.

Question 3 (optional)

Recall the example from the lecture about the proportion of patients attending a diabetes clinic who respond to lifestyle changes. How would the likelihood have changed if 50 patients had attended the clinic that day and been included in the clinician's investigation, rather than 20?

(a) Plot the likelihood function for the proportion π of responsive patients, if the clinician had included 50 patients in her study and observed that 32 of them were responsive.

(b) Calculate the MLE for π based on these alternative data.

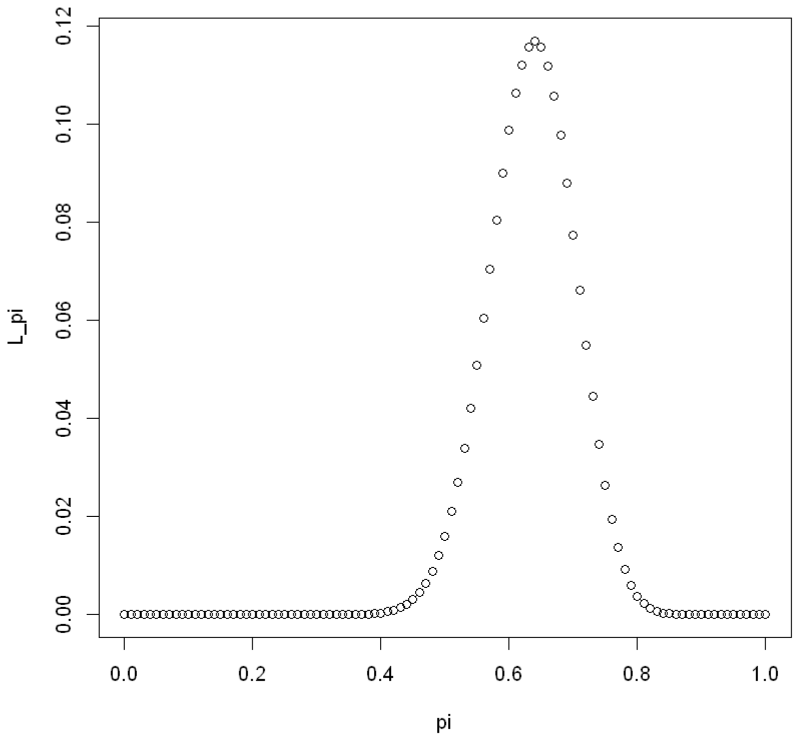
Solution: From the lecture, $\hat{\pi} = \frac{x}{n}$*, so* $\hat{\pi} = \frac{32}{50} = 0.64$

(c) The table below shows the number of responders out of five different samples of patients, of different sizes. By changing the values in your code for part *a*, plot the likelihood for each sample. How does the likelihood change as the size increases?

Patients (n)	Responders (x)
10	6
20	11
50	32
100	58
500	328

In [1]:

```
# Range of possible values to plot
pi = seq(0,1,by = 0.01)

# Data
n = 50
x = 32

L_pi <- dbinom(x = x, size = n, p = pi)

plot(x = pi, y = L_pi)
```



The code below allows you to add the different likelihoods to the same plot.

In [2]:

```
# This code sets up a colourblind-friendly palette to use in the plot
pal <- RColorBrewer::brewer.pal(5, "Dark2")

# Calculate and plot the likelihood for each sample, for the same values of pi
L_pi_10 <- dbinom(x = 6, size = 10, p = pi)
plot(x = pi,
     y = L_pi_10,
     col = pal[1],           # point colour
     pch = 19,               # plotting character
     ylab = "Likelihood")    # axis label

L_pi_20 <- dbinom(x = 11, size = 20, p = pi)
points(x = pi, y = L_pi_20, col = pal[2], pch = 19)

L_pi_50 <- dbinom(x = 32, size = 50, p = pi)
points(x = pi, y = L_pi_50, col = pal[3], pch = 19)

L_pi_100 <- dbinom(x = 58, size = 100, p = pi)
points(x = pi, y = L_pi_100, col = pal[4], pch = 19)

L_pi_500 <- dbinom(x = 328, size = 500, p = pi)
points(x = pi, y = L_pi_500, col = pal[5], pch = 19)

# Add a legend for the colours
legend(x = 0.1, y = 0.25, col = pal, pch = 19, legend = c(10,20,50,100,500), title = "S
ize of sample")
```

```
Error in RColorBrewer::brewer.pal(length(patients), "Dark2"): object 'pati
ents' not found
Traceback:

1. RColorBrewer::brewer.pal(length(patients), "Dark2")
```

In [ ]: