

Homework #1

Chapters 1.1, 1.2 and 1.3

Brynjólfur Gauti Guðrúnar Jónsson

1. Sample survey

Suppose we are going to sample 100 individuals from a county (with population size much larger than 100) and ask each sampled person whether they support policy Z or not. Let $Y_i = 1$ if person i in the sample supports the policy, and $Y_i = 0$ otherwise.

- Assume Y_1, \dots, Y_{100} are, conditional on θ , i.i.d. binary random variables with expectation θ . Write down the joint distribution of $\Pr(Y_1 = y_1, \dots, Y_{100} = y_{100})$ in a compact form. Also write down the form of $\Pr(\sum_{i=1}^{100} Y_i = y)$.
- For the moment, suppose you believed that $\theta \in \{0, 0.1, \dots, 0.9, 1.0\}$. Given that the results of the survey were $\sum_{i=1}^{100} Y_i = 73$, compute $\Pr(\sum_{i=1}^{100} Y_i = 73 | \theta)$ for each of these 11 values of θ and plot these probabilities as a function of θ (point mass at each value of θ).

```
```{r}
Include your code here
```
```

- Now suppose you originally had no prior information to believe one of these θ -values over another, and thus $\Pr(\theta = 0.0) = \Pr(\theta = 0.1) = \dots = \Pr(\theta = 0.9) = \Pr(\theta = 1.0) = \frac{1}{11}$. Use Bayes' rule to compute $p(\theta | \sum_{i=1}^{100} Y_i = 73)$ for each θ -value. Make a plot of this posterior distribution as a function of θ (point mass at each value of θ).

```
```{r}
Include your code here
```
```

- Now suppose you allow θ to be any value in the interval $[0, 1]$. Using the uniform prior density for θ , namely, $p(\theta) = 1$, derive and plot the posterior density of θ as a function of θ . According to the posterior density, what is the probability of $\theta > 0.8$?

```

```{r}
Include your code here
```

```

- e. Why are the heights of posterior densities in *c.* and *d.* not the same?

2. Random numbers, probability density functions (pdf) and cumulative density functions (cdf)

The goal of this exercise is to generate random numbers, plot the histogram, the empirical pdf and cdf for these numbers, and see how they compare to the theoretical pdf and cdf. The goal is also to compare the sample mean and standard deviation to the theoretical mean and standard deviation.

- a. Generate $B = 3000$ numbers from the gamma distribution with parameters $\alpha = 2$ and $\beta = 0.1$. Compute the sample mean and the sample standard deviation and compare to the theoretical mean and standard deviation.

```

```{r}
Include your code here
```

```

- b. Plot the theoretical density (pdf) of the gamma distribution. Plot the empirical density based on the data on the same graph. Plot the histogram of the data on another graph.

```

```{r}
Include your code here
```

```

- c. Plot the theoretical cumulative density function (cdf) of the gamma distribution. Plot the empirical cumulative density based on the data on the same graph.

```

```{r}
Include your code here
```

```