10.

Intro

After some fast search on Google about the tramcar problem, I discoreved that this is actually a reasonable famous problem in statistics. However, it seems that the best solution, as one can imagine, does not use a continuous distribution. My guess is that the author wanted us to use the continuous distribution approximation to see how to use the results of the Bayesian analysis of the last exercise in a real problem. Fortunately, despite having five itens, this problem isn't very long.

Solution

a)

Using the posterior distribution derived in Exercise 9:

$$p(\theta|D) = Pareto(\theta|N+K, max(m,b)) = Pareto(\theta|1+0, max(100,0)) = Pareto(\theta|1,100) \tag{1}$$

b)

The mean of $Pareto(\theta|k,m)$ is equal to $E[\theta|D] = \frac{km}{k-1}, k > 1$. Since our distribution has $k = 1$, the mean doesn't exist. The mode of the distribution is given by $mode(\theta|D) = m = 100$ The median of the distribution is given by $median(\theta|D) = m(2)^{\frac{1}{k}} = 100(2)^{\frac{1}{1}} = 200$

c)

Let's compute the predictive density over the next taxi number. First, we need to compute the posterior $p(\theta|D)$:

$$p(\theta|D) = Pareto(\theta|N+K, max(m,b)) = Pareto(\theta|1+0, max(m,0)) = Pareto(\theta|1,m) \tag{2}$$

Since the posterior is a Pareto distribution like the prior, we can use it as a "prior" for the inference on D' and use the results from Equations 3.93 - 3.95, as the author suggest. So, our new "prior" has the follwing distribution $p(\theta|D) = Pareto(\theta, K' = 1, b' = m)$. The number of samples is $N' = 1$ and $m' = max(D') = x$. Now, we can calculate the predictive distribution using 3.95.

$$\begin{aligned}p(x|D,\alpha) = \frac{K'}{(N'+K')b'^{N'}} I(x \leq m) + \frac{K'b'^{K'}}{(N'+K')m'^{N'+K'}} I(x > m) = \\ \frac{1}{2m} I(x \leq m) + \frac{m}{2x^2} I(x > m)\end{aligned} \tag{3}$$

d)

$$\begin{aligned}p(x = 100|D,\alpha) &= \frac{1}{2m}\mathcal{I}(100 \leq m) + \frac{m}{20000}\mathcal{I}(100 > m) \\ p(x = 50|D,\alpha) &= \frac{1}{2m}\mathcal{I}(50 \leq m) + \frac{m}{5000}\mathcal{I}(50 > m) \\ p(x = 150|D,\alpha) &= \frac{1}{2m}\mathcal{I}(150 \leq m) + \frac{m}{45000}\mathcal{I}(150 > m)\end{aligned} \tag{4}$$

e)

To be completly honest, I don't know the right answer to give here. Some suggestion that I think may help the model are: work with a discrete probability distribution, since our data is discrete. Another useful suggestion would be the use of an informative prior, since making a reasonable estimate of the numbers of taxi in a city does not require expert knowledge.

Conclusion

In this question we used the results of the Bayesian analysis of the uniform distribution infered in the last question in a real problem (the tramcar problem). In item a, we saw how to compute the posterior, given one observation. The posterior $p(\theta|D) = Pareto(\theta|1, 100)$ tell us that the parameter must be equal or greate than 100. Also, since we only made one observation, there is a good chance that the parameters may not be that close to 100.

In item b, we calculated some numbers that give useful insight on the distribution (the mean, the mode and the median). We discoreved that there is no mean for the posterior. This is very important, because it tells us that if we work with the sample average, there is no meaning in increase the number of samples, hoping to arrive at a better estimate of the true mean. We also discoreved that the mode is equal 100. This is straightfoward, since the distribution is a decreasing function, starting at 100. Finally, the median is located at 200. This corroborate our statement that the true parameter can be distant from the maximum value of our dataset (m = 100).

In item c, we computed the predictive density over the next taxcab number x, given a previous observation $D = \{m\}$. We discovered that the distribution is uniform if $x < m$ and decreases at a rate of $\frac{1}{x^2}$ otherwise.

In item d, we just used the result of c, to compute some particular examples. Since we don't have the value of m, we cannot give a actual number for each result.

In item e, we had to make some suggestion to improve the model. As I said, I am not sure if those two recommendations are the "right" answer, because I don't remember anything resembling this problem being mentioned in the text. Nevertheless, both suggestion are good paths to improve the results.