Physalia Bayesian statistics course (Shravan Vasishth/Bruno Nicenboim)

Statistics Quiz 2019-03-11

1. (a) .
2. (a) .

   (b) .
3. (a) .

   (b) .

   (c) .

   (d) .

   (e) .
4. (a) .

   (b) .

   (c) .
5. (a) .

   (b) .

   (c) .

6. (a) | | | | | | . | | |

   (b) | | | | | | . | | |

7. (a) | | | | | | . | | |

   (b) | | | | | | . | | |

   (c) | | | | | | . | | |

   (d) | | | | | | . | | |

8. (a) | | | | | | . | | |

   (b) | | | | | | . | | |

   (c) | | | | | | . | | |

   (d) | | | | | | . | | |

1. [**Please give your answer as a number with three decimal places. Example: 0.010.**]

   Given a normal distribution with mean 52 and standard deviation 103, use the pnorm function to calculate:

   (a) the probability of obtaining values between 81 and 75 from this distribution.

2. Consider a normal distribution with mean 1 and standard deviation 1.

   Compute, to three decimal places, the lower and upper boundaries such that:

   (a) the area (the probability) to the left of the lower boundary is 0.07

   (b) the area (the probability) to the left of the upper boundary is 0.78.

3. [Give answers up to three decimal places for each case.]

   Take an independent random sample of size 145 from a normal distribution with mean 178, and standard deviation 64. Next, we are going to pretend we don't know the population parameters (the mean and standard deviation). We compute the MLEs of the mean and standard deviation using the data and get the sample mean 182.279 and the sample standard deviation 63.37. Compute:

   (a) the estimated standard error using the sample standard deviation provided above.

   (b) What are your degrees of freedom for the relevant t-distribution?

   (c) Calculate the **absolute** critical t-value for a 95% confidence interval using the relevant degrees of freedom you just wrote above.

   (d) Next, compute the lower bound of the 95% confidence interval using the estimated standard error and the critical t-value.

   (e) Finally, compute the upper bound of the 95% confidence interval using the estimated standard error and the critical t-value.

4. [Give answers up to three decimal places for each case. Example: 0.123.]

   Calculate the following probabilities:

   Given a normal distribution with mean 52 and standard deviation 2, what is the probability of getting

   (a) a score of 43 or less

   (b) a score of 43 or more

   (c) a score of 52 or more

5. [Give answers up to three decimal places for each case. Example: 0.123.]

   Given a normal distribution with mean 52 and standard deviation 10, what is the probability of getting

   (a) a score of 47 or less
   (b) a score between 49 and 55
   (c) a score of 53 or more

6. Given a normal distribution with mean 56.446 and standard deviation 0.986. There exist two quantiles, the lower quantile q1 and the upper quantile q2, that are equidistant from the mean 56.446, such that the area under the curve of the Normal probability between q1 and q2 is 85%. Find q1 and q2.

   Give your answer to three decimal places.

   (a) lower bound:
   (b) upper bound:

7. [**Please give your answer as a number with three decimal places. Example: 0.010.**]

   Given the data point 18.395. The function dnorm gives the likelihood given a data point (or multiple data points) and a value for the mean and the standard deviation (sd). Using dnorm, compute

   (a) the likelihood of the data point 18.395 assuming a mean of 12 and standard deviation 5.
   (b) the likelihood of the data point 18.395 assuming a mean of 11 and standard deviation 5.
   (c) the likelihood of the data point 18.395 assuming a mean of 10 and standard deviation 5.
   (d) the likelihood of the data point 18.395 assuming a mean of 9 and standard deviation 5.

8. [**Please give each answer as a number with three decimal places. Example: 0.010.**]

   You are given 10 independent and identically distributed data points that are assumed to come from a Normal distribution with unknown mean and unknown standard deviation:

```
> x

 [1] 501 505 484 491 489 504 488 492 509 501
```

The function dnorm gives the likelihood given multiple data points and a value for the mean and the standard deviation (sd). The log-likelihood can be computed by typing dnorm(...,log=TRUE).

The product of the likelihoods for two independent data points can be computed like this: Suppose we have two independent and identically distributed data points 5 and 10. Then, assuming that the Normal distribution they come from have mean 10 and sd 2, the joint likelihood of these is:

```
> dnorm(5,mean=10,sd=2)*dnorm(10,mean=10,sd=2)
```

```
[1] 0.001748195
```

It is easier to do this on the log scale, because then one can add instead of multiplying. This is because $\log(x \times y) = \log(x) + \log(y)$. For example:

```
> log(2*3)
```

```
[1] 1.791759
```

```
> log(2) + log(3)
```

```
[1] 1.791759
```

So the joint log likelihood of the two data points is:

```
> dnorm(5,mean=10,sd=2,log=TRUE)+dnorm(10,mean=10,sd=2,log=TRUE)
```

```
[1] -6.349171
```

Even more compactly:

```
> sum(dnorm(c(5,10),mean=10,sd=2,log=TRUE))
```

```
[1] -6.349171
```

Compute the following quantities:

(a) Given the 10 data points above, calculate the maximum likelihood estimate (MLE) of the expectation.

(b) The sum of the log-likelihoods of the data-points x, using as the mean the MLE from the sample, and standard deviation 5.

(c) What is the sum of the log-likelihood if the mean used to compute the log-likelihood is 494.4?

(d) Which value for the mean, the MLE or 494.4, gives the higher log-likelihood? As your answer, write either the MLE or 494.4.