LOT 2020 Bayesian statistics course (Shravan Vasishth/Bruno Nicenboim)

Statistics Quiz 2020-01-12

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 49 and standard deviation 98, use the pnorm function to calculate:

   (a) the probability of obtaining values between 298 and 192 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 0.38

   (b) the area (the probability) to the left of the upper boundary is 0.59.

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

   Take an independent random sample of size 169 from a normal distribution with mean 163, and standard deviation 71. 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 151.277 and the sample standard deviation 70.646. 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 55 and standard deviation 3, what is the probability of getting

   (a) a score of 43 or less

   (b) a score of 43 or more

   (c) a score of 56 or more

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

   Given a normal distribution with mean 49 and standard deviation 6, what is the probability of getting

   (a) a score of 44 or less
   (b) a score between 46 and 52
   (c) a score of 50 or more

6. Given a normal distribution with mean 58.821 and standard deviation 0.575. There exist two quantiles, the lower quantile q1 and the upper quantile q2, that are equidistant from the mean 58.821, such that the area under the curve of the Normal probability between q1 and q2 is 80%. 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 13.062. 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 13.062 assuming a mean of 12 and standard deviation 5.
   (b) the likelihood of the data point 13.062 assuming a mean of 11 and standard deviation 5.
   (c) the likelihood of the data point 13.062 assuming a mean of 10 and standard deviation 5.
   (d) the likelihood of the data point 13.062 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] 499 491 490 519 486 507 493 498 516 503
```

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 has mean 10 and sd 2, the joint likelihood of these is:

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

[1] 0.0017482
```

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] 1.7918

> log(2) + log(3)

[1] 1.7918
```

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] -6.3492
```

Even more compactly:

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

[1] -6.3492
```

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 498.2?

(d) Which value for the mean, the MLE or 498.2, gives the higher log-likelihood? As your answer, write either the MLE (the actual number, not the words MLE!) or 498.2.