Juliette Woodrow
CS 109

Section #4
Feb 4, 2026

Section 4

1 Mental Gymnastics

Odysseus is preparing for the ancient Greek Olympic games, representing the Greek island of Ithaca. He has decided to participate in the Men's Gymnastics event and is aiming to set a new high score for the event.

It is common knowledge that the Olympic judges fairly score athletes according to their gymnastic skill level. Since Greek athletes' skill levels are normally distributed, athlete scores follow a normal distribution, with an average of 8.0 and a variance of 0.6. As one of these many Greek athletes, Odysseus knows that his score will follow this trend.

Odysseus feels unsure about whether he can accomplish his goal, so he asks his friend Mentor[1] for advice. Mentor watches Odysseus perform his gymnastic routine, full of tricks and twists, and remarks: "Were you to face off against a hundred athletes, I expect the judges would find you match or surpass the very best of them!"

1. Based on Mentor's assessment of Odysseus' performance, what is the minimum score Odysseus' routine would earn if he were to be judged in the Olympics? In other words, what is the minimum score Odysseus will achieve, given that he scores higher than 99% of athletes?

Let X = the score of a competitor at the gymnastic games. We know that $X \sim \mathcal{N}(\mu = 8.0, \sigma^2 = 0.6)$. Furthermore, we know that Mentor thinks that Odysseus scores higher than 99% of other competitors. Hence, his score must be at the 99th percentile of X or above.

In order to calculate the value that corresponds to the 99th percentile of X, we need only use the inverse Phi formula Φ^{-1}. This formula converts an inputted probability p to the value of the standard normal distribution v for which $P(X \leq v) = p$ (i.e. the value for which the outcome of the random variable is less than that value with exactly p probability). Then, we can "un-standardize" the result of the inverse Phi function to get the minimum score that Odysseus's performance would achieve! Mathematically, assuming v is Odysseus' score, we see that:

$$\Phi(\frac{v - \mu}{\sigma}) = p$$

$$\Phi(\frac{v - 8.0}{\sqrt{0.6}}) = 0.99$$

[1] Fun fact: The modern English word "mentor" comes from the name of Telemachus' advisor, Mentor! Telemachus is Odysseus' son in the Ancient Greek Epic *The Odyssey*. As you might guess, the modern word "odyssey" also comes from Odysseus' epic journey in *The Odyssey*.

Thus, solving for v using the inverse Phi function, we get:

$$\Phi(\frac{v-\mu}{\sigma}) = p$$

$$\frac{v-\mu}{\sigma} = \Phi^{-1}(p)$$

$$v - \mu = \Phi^{-1}(p) * \sigma$$

$$v = \Phi^{-1}(p) * \sigma + \mu$$

$$v = \Phi^{-1}(0.99) * \sqrt{0.6} + 8.0 \approx 9.80198$$

2. The current high score for men's gymnastics is 9.9. Given Odysseus will receive at least his minimum score from part 1, how likely is it that Odysseus is able to set a new high score for men's gymnastics at the Olympics?

We know from part 1 that Odysseus' score is more than approximately 9.8. Thus, we must figure out how likely it is that Odysseus' score is greater than 9.9, given it's greater than 9.8.

Let X = the score of a participant at the Olympics. We must calculate $P(X > 9.9 | X > 9.8)$.

We can decompose this probability into $\frac{P(X>9.9, X>9.8)}{P(X>9.8)}$.

A useful observation is that $X > 9.9$ and $X > 9.8$ if, and only if, $X > 9.9$. Thus, we can make the substitution $P(X > 9.9, X > 9.8) = P(X > 9.9)$.

Now, in order to calculate $P(X > x)$, we can use the complement along with the normal CDF, since $P(X > x) = 1 - P(X \leq x) = 1 - \Phi(\frac{x-\mu}{\sigma})$.

Thus, we have:

$$P(X > 9.9 | X > 9.8) = \frac{P(X > 9.9, X > 9.8)}{P(X > 9.8)} = \frac{P(X > 9.9)}{P(X > 9.8)}$$

$$= \frac{1 - P(X \leq 9.9)}{1 - P(X \leq 9.8)} = \frac{1 - \Phi(\frac{9.9-8.0}{\sqrt{0.6}})}{1 - \Phi(\frac{9.8-8.0}{\sqrt{0.6}})}$$

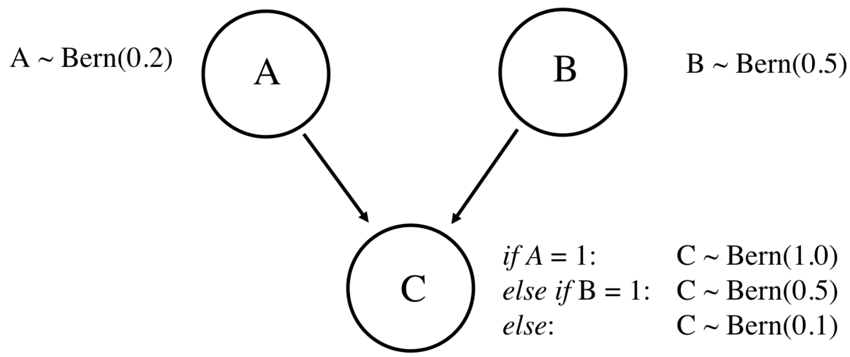
$$= \frac{1 - 0.9929}{1 - 0.99} \approx 0.71$$

So, Odysseus has a good shot at breaking the record, not to mention bringing home the gold!

2 Understanding Bayes Nets

	A = 0		**A = 1**	
	B = 0	B = 1	B = 0	B = 1
C = 0	0.36	0.20	0.00	0.00
C = 1	0.04	0.20	0.10	0.10

The **joint probability table (above)** for random variables A, B and C is equivalent to the **bayesian network (below)**. Both give the probability of any combination of the random variables. In the Bayes network the probability of each random variable is provided given its causal parents.



a. Use the bayesian network to explain why $P(A = 0, B = 1, C = 1) = 0.20$.

> We can intentionally break down this joint probability using the Chain Rule, so that we are left with probabilities of events that the bayes net directly gives us: $P(A = 0, B = 1, C = 1) = P(A = 0)P(B = 1)P(C = 1|A = 0, B = 1) = 0.8 * 0.5 * 0.5 = 0.2$.

b. What is $P(A = 1|C = 1)$? Use the information in the table to justify your answer.

Using the table, we see that

$$P(A = 1|C = 1) = \frac{0.1 + 0.1}{0.1 + 0.1 + 0.2 + 0.04} = \frac{0.2}{0.44} = \frac{5}{11}$$

If we wanted to find this probability using the bayes net instead, it would be possible, but it would require using Bayes' Theorem, since the structure of the network makes it much easier to think about $P(C = 1|A = 1)$ than $P(A = 1|C = 1)$.

c. Is A independent of B? Explain your answer.

Yes. This follows directly from the structure of the bayesian network, because A and B have no shared ancestors. Alternatively, note that from the table or the network, we can calculate that $P(A = a, B = b) = P(A = a)P(B = b)$, which satisfies the definition of independence.

d. Is A independent of B **given** $C = 1$? Explain your answer.

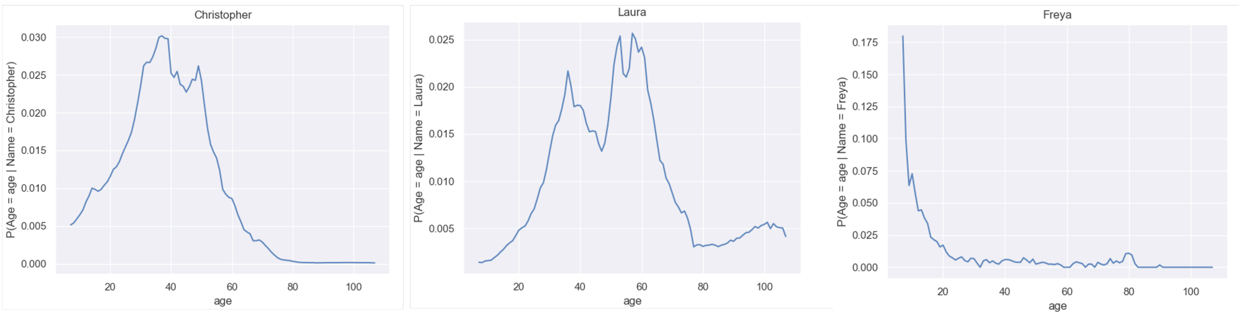
No. From the table, we can see that $P(B = 1|A = 0, C = 1) = \frac{0.2}{0.04+0.2} \neq P(B = 1|A = 1, C = 1) = \frac{0.1}{0.1+0.1}$. So given $C = 1$, knowing the value of A informs us about the value of B, and therefore A and B are not conditionally independent given C.

This probably feels counterintuitive! Here's an example of what these random variables could be to illustrate how this is possible. Let A, B, and C each be Bernoulli RVs as in the bayes net. C is 1 if you are holding a fresh mandarin, 0 otherwise. A is 1 only if you recently received a mandarin from Chris in lecture, and B is 1 only if you recently bought mandarins at the store. If you have a mandarin ($C = 1$), it's very likely that you got it either from Chris ($A = 1$) or from the store ($B = 1$). So then, if you have a mandarin and tell me you haven't been to the store recently ($B = 0$), that leaks information about A, because I can guess that you probably received your mandarin from Chris!

This phenomenon is sometimes called 'Explaining Away' if you're curious to read more.

3 Name2Age Inference

What is the probability distribution of someone's age given just their name? Here are a few example for the names 'Christopher' 'Laura' and 'Freya':



The U.S. Government released a dataset of the frequencies, by year, of all given names recorded in U.S. births at least 5 times. You can access this data via the function `get_count(name, year)` which returns the number of babies named `name` born in `year`. Since this data provides the joint distribution, it can be used to solve inference problems.

Write a function in pseudocode that 1) takes in a name and infers the conditional distribution $P(\text{Age} = \text{age} | \text{Name} = \text{name})$ across all of the ages covered by the dataset, and 2) plots this conditional probability function (see the plots above as examples).

```
def run_name_query(name, years_list):
```

For the solution, see this chapter in the course reader: `https://probabilityforcs.firebaseapp.com/book/name2age`
The code (including the solution in Python) and data are also available here: `http://web.stanford.edu/class/cs109/section/5/babynames.zip`