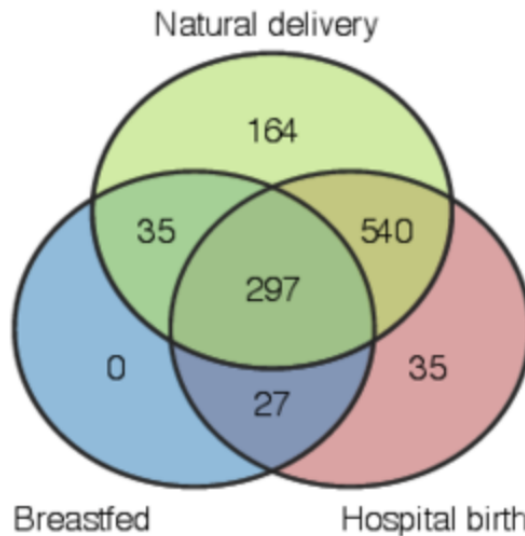


1. (5 marks) **Birth.** As part of an analysis of gut microbiome variation Dutch researchers collected data on early-life factors. They distinguished between people born in a hospital vs born at home, between natural delivery and caesarian section, and between being breastfed or not. Here's a summary of their counts across these categories:



- What is the probability that someone in their dataset was born in a hospital?
 - Did a hospital birth make one more or less likely to be breastfed?
 - What combination of factors made one most likely to be breastfed?
 - What is the interpretation of the 0 entry in the diagram?
2. (10 marks) **Our relatives.** Chimpanzees and bonobos are humans' closest relatives among the great apes, and both have genomes that overlap a great deal with ours.

“Bonobos and chimpanzees are highly similar to each other in many respects. However, the behaviour of the two species differs in important ways. For example, male chimpanzees use aggression to compete for dominance rank and obtain sex, and they cooperate to defend their home range and attack other groups. By contrast, bonobo males are commonly subordinate to females and do not compete intensely for dominance rank. They do not form alliances with one another and there is no evidence of lethal aggression between groups. Compared with chimpanzees, bonobos are playful throughout their lives and show intense sexual behaviour that serves non-conceptive functions and often

involves same-sex partners. Thus, chimpanzees and bonobos each possess certain characteristics that are more similar to human traits than they are to one another's."¹

Genomes can be compared according to the amount of shared genomic material measured in Megabases, Mb². Human and chimpanzee genomes share around 69 Mb of genome segments, chimpanzees and bonobos around 85.4 Mb, and bonobos and humans around 68.7 Mb. Of their shared segments chimpanzees and bonobos have 17.2 Mb that are not shared with humans. In total chimpanzees have slightly more genomic segments than bonobos, 87 Mb vs 86.6 Mb. Across the three species there are 102.8 Mb unique segments.

- a. What is the probability that a segment is common across all three species?
- b. Given a segment is shared, which pair of species is it most likely to be shared between? With what probability?
- c. Given that a segment is shared between humans and another species, which other species is it more likely to be? With what probability?

3. (10 marks) **Bird Nests.** In a study of bird nesting habits Martin, Aitken and Wiebe³ reported the following observations about their study plot:

- It contained 10,719 available trees, 690 of which contained at least one nest.
 - It contained four tree species: 1,562 Quaking aspen, 3,260 Douglas-fir, 4,296 Lodgepole pine, and 1,601 Hybrid spruce.
 - There were nests in 658 of the Quaking aspen, 9 of the Douglas-fir, 18 of the Lodgepole pine, and 6 of the Hybrid spruce.
 - 1,612 nests were observed in Quaking aspen trees, 19 in Douglas-fir, 48 in Lodgepole pine and 10 in Hybrid spruce.
 - In one subsample of trees with nests 87 trees had a single nest, 51 of which were in use that season, 99 trees had 2 or 3 nests, in 33 of which at least one nest was in use, and 33 trees had more than 3 nests, only 5 of which had nests in use that season.
- a. If a tree is selected at random what is the probability that it is a Quaking aspen? That it is a Douglas-fir?
 - b. If a tree is selected at random what is the probability that it has at least one nest in it?
 - c. If a tree is a Quaking aspen what is the probability that it has at least one nest in it? What if it is a Douglas-fir?

¹ The bonobo genome compared with the chimpanzee and human genomes. Kay Prüfer et. al. Nature 486, 527–531(2012).

² Megabase: unit of length for DNA fragments equal to 1 million nucleotides.

³ Martin, K., K.E.H. Aitken and K.L. Wiebe. 2004. Nest sites and nest webs for cavity-nesting communities in interior British Columbia, Canada: nest characteristics and niche partitioning. Condor 106: 5-19

- d. If a tree has a nest in it what is the probability that it is a Quaking aspen? That it is a Douglas-fir?
 - e. If you spot a nest in a tree what is the probability that it is the only nest in the tree?
 - f. If you spot a nest in a tree what is the probability that it has been in use this season?
4. (20 marks) *March Madness*. Each Spring⁴ the NCAA holds its championship basketball tournament. This is a knockout tournament (teams are eliminated after their first loss) in which 64 teams compete to determine the champion. A selection committee selects the teams to participate and *seeds* them in four regions (South, West, Midwest and East) according to their estimated strength from strongest (1) to weakest (16).



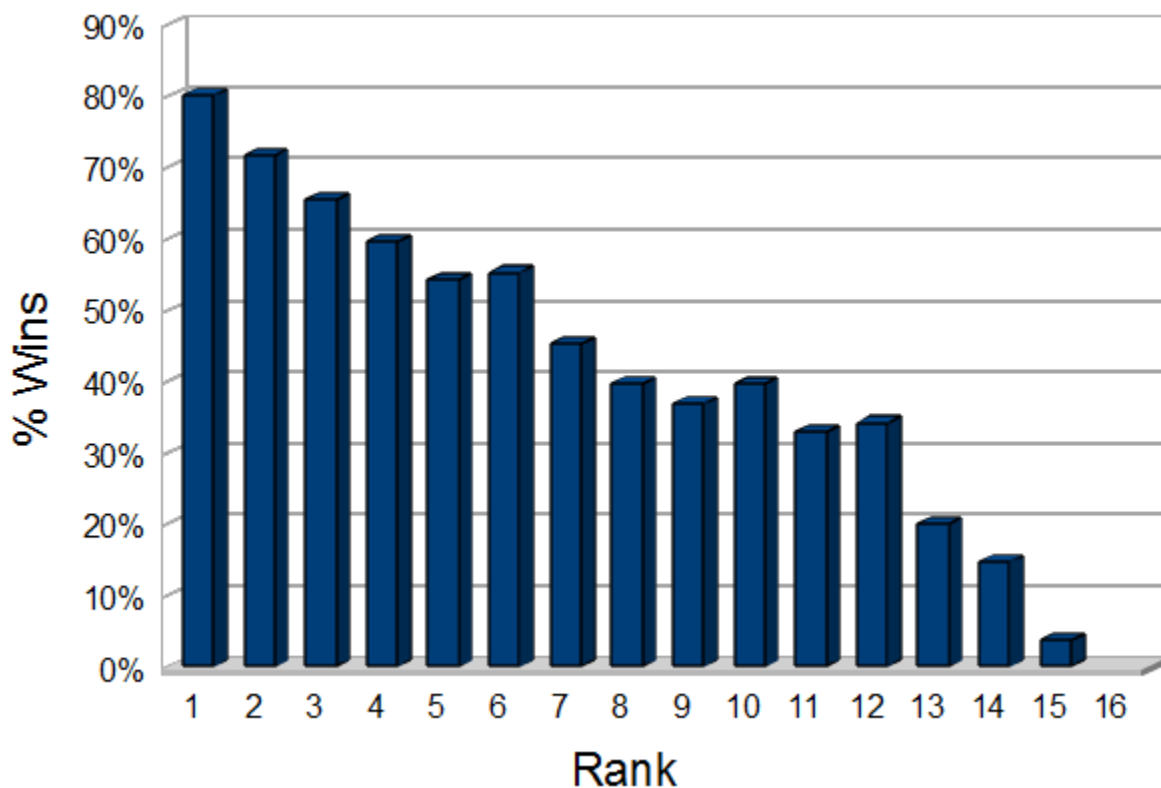
Each year there are contests to try and predict the outcomes of the games, and in particular to pick the outcomes of *all* the games correctly. One year the contest that caught everyone's attention was the Quicken Loans Billion Dollar Bracket Challenge which offered to award

⁴ Well until 2020 when COVID-19 forced cancellation of the tournament!

\$1,000,000,000 (a billion dollars!) to any person who could correctly pick the winners of **all** the games, and \$100,000 each to the people with the 20 most accurate “imperfect” brackets.

- How many games will be played in the tournament?
- How many possible ways are there for the tournament to play out?
- If one just guessed randomly at the outcomes of each game what is the probability of getting them all correct?

The odds in c. above are poor. How can we do better? By making informed guesses rather than random ones! The outcomes of all games are not of course equally likely. For example as of 2018 no number 1 seed had ever (in 116 games!) lost to a number 16 seed. Here are the %s of their games teams have historically won according to their seeding rank.



- Using the data above what would you predict the outcomes of the games in the tournament to be?
- What is the probability that your prediction in d. is correct?
- You may have found that the data in the graph above were not helpful in predicting the outcomes of some games in d. More detailed data is available at

<http://rotoguru2.com/hoop/mad/seedw1.html> Does it enable you to make a better prediction? If so, what is your improved prediction and how good do you think it is? If not, why not?

- g. If the contest were to be held next year do you think someone would win it? Why or why not?
5. (5 marks) *Pitcher problems.* You are at the side of a river. You have a three-litre pitcher and a seven-litre pitcher. The pitchers do not have markings to allow measuring smaller quantities—all you can do is fill them from the river, dump them into the river or pour from one to the other.
- a. You need one litre of water. How can you measure one litre?
- b. You need two litres of water. Can you measure two litres?