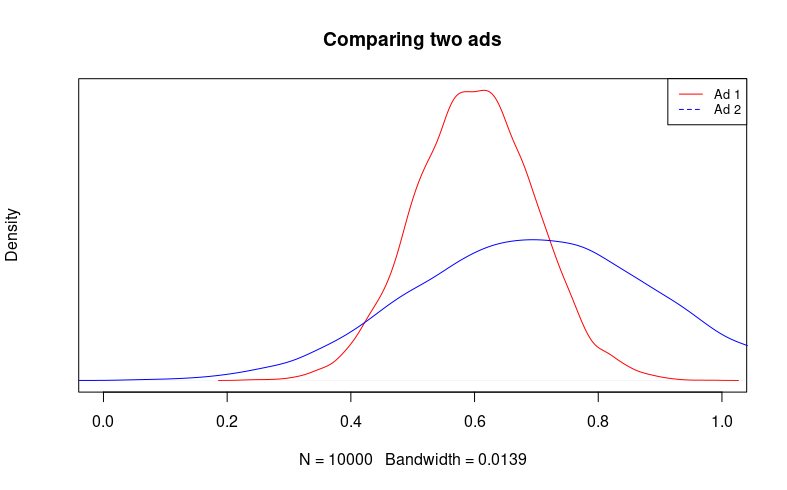
**Week 4 Lab: Comparing Distributions**

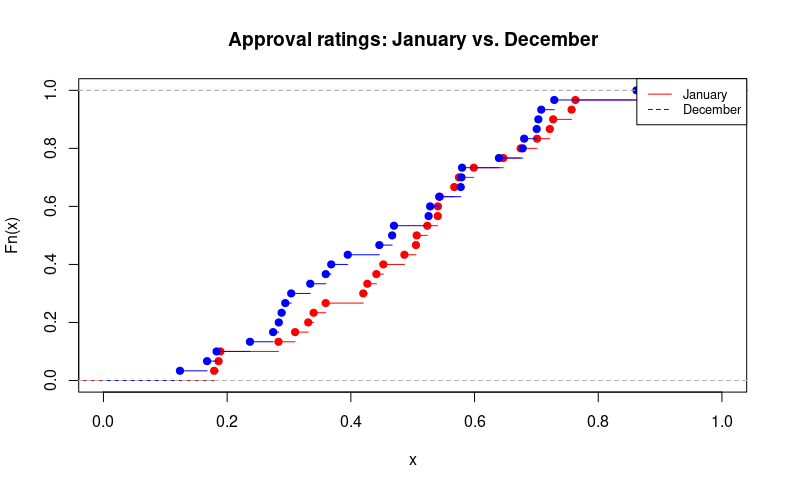
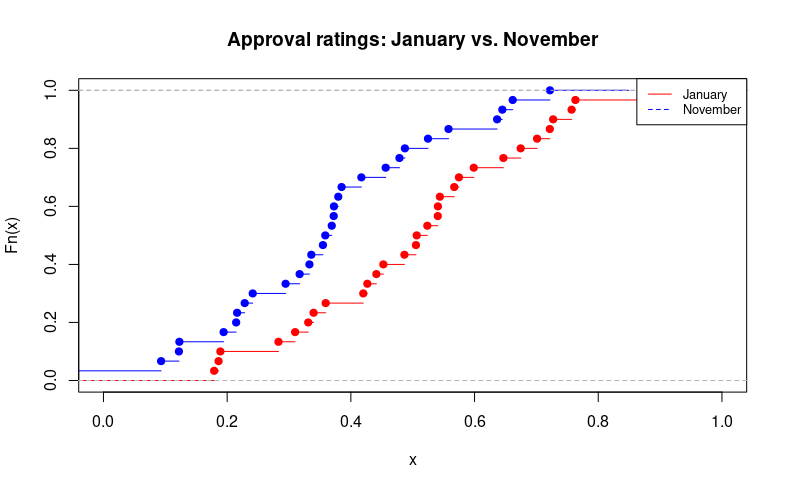
1. You are running a polling firm for a candidate in political office. You are trying to decide which ad a candidate wants to run will best increase their chance of winning office. You compare two different ads, show each to 100 people, and then ask how likely each person is to vote for them. After running this on many different groups, you generate two probability distributions: one for support after seeing Ad 1, the other for support after seeing Ad 2. The x-axis is the proportion of respondents who supported the ad. Answer the following question through a visual comparison

a) Describe the distributions for Ad 1 and Ad 2.

b) Which ad has a higher expected value? Which ad has a greater variance?

c) The candidate needs only above 50% of the vote to win the race. Which ad would you recommend that they use? Why? Your answer should discuss the various scenarios that might happen and their likelihood under both Ad 1 and Ad 2.

2. A politician has commissioned a polling agency to provide them with approval ratings on a monthly basis. They take one poll per day (30 per month) and then present the evidence as a CDF, where the x-axis conveys the approval rating, and the y-axis is the percentile range of each value. They are wondering whether they have been gaining support recently, so you analyze two charts: the CDF of January vs. December, and the CDF of January vs. November.



a) Rank the three months presented, January, December, and November by the expected value of the approval rating. Does the variance change between the three months?

b) Remember that random variables are uncertain, and will have natural fluctuations even if the data-generating process of that variable has not changed. Does it seem that there has been a meaningful shift in approval rating since December? How about since November? Please explain your answer.