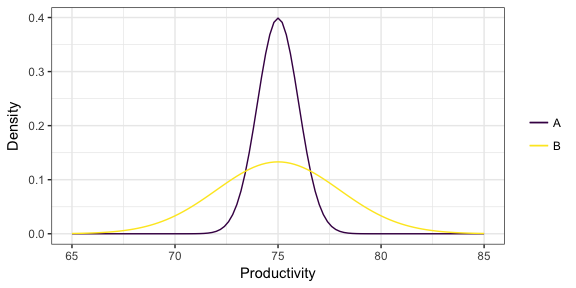
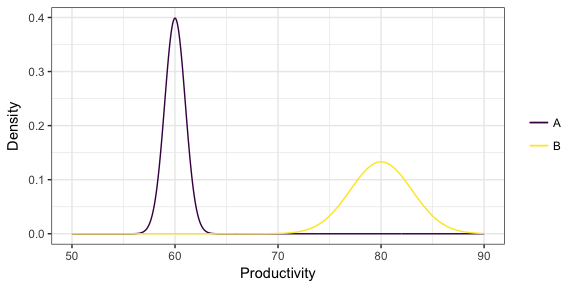
1. **Both are predicted to have the same productivity score of 75 but have the following probabilistic forecasts.**

[](https://github.com/vincenzocoia/BAIT509/blob/master/class_meetings/cm08-beyond_mean_mode_files/figure-html/unnamed-chunk-17-1.png)

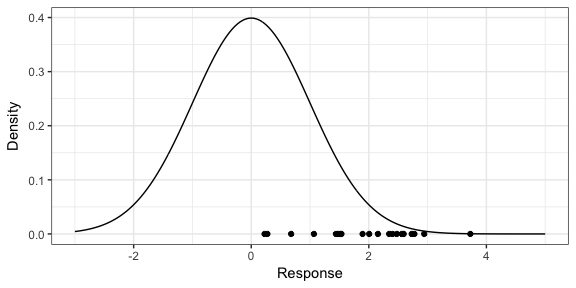
* I would choose candidate A. The pros of choosing the candidate include the high probability that he will perform to the expected productivity level of 75 but cons include that we are restricting our self to a lower bandwidth of productivity, since the candidate B has chances of productivity higher than 75 up to 85 and candidate A has only till 77.

1. **Two "non-overlapping" forecasts:**

[](https://github.com/vincenzocoia/BAIT509/blob/master/class_meetings/cm08-beyond_mean_mode_files/figure-html/unnamed-chunk-18-1.png)

* I would choose candidate B. The pros of choosing this candidate include higher chances of high productivity in comparison to candidate A, but cons include higher bandwidth and lower probability density that for sure the candidate will perform to the expected productivity level. B has higher productivity but less chances of expected and more variance and candidate A has very high chance of lower productivity but less variance.

1. **You've formed a probabilistic forecast for a particular value of the predictors, displayed below as a density. You then collect test data for that same value of the predictor, indicated as the points below the density. What is the problem with the probabilistic forecast?**

[](https://github.com/vincenzocoia/BAIT509/blob/master/class_meetings/cm08-beyond_mean_mode_files/figure-html/unnamed-chunk-19-1.png)

* From the test set validation, we can observe that our actual data is always towards the right of our probabilistic forecast.
* The forecast is not normally distributed with the expected response and is highly skewed.
* The center of the forecast is not the expected outcome of the response variable.