0.632.

Say, we are doing bootstrap.

Let our original sample be $x_1, x_2, \ldots, x_n$.

With bootstrap, we draw with replacement from the original sample.

Focusing on, say x_1, the probability of it not being chosen in the draw is

$$1 - \frac{1}{n}$$

But, we have n independent draws.
So, the total probability of never choosing x_1 is

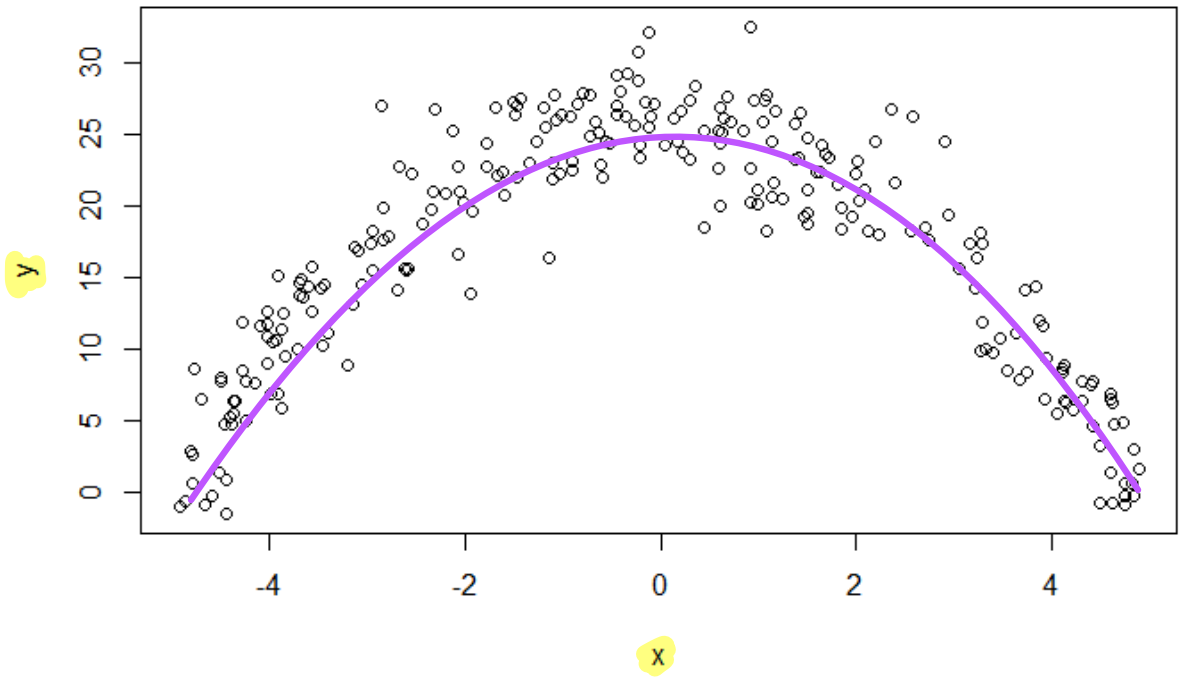
$$\left(1 - \frac{1}{n}\right)^n \xrightarrow[n \to \infty]{} e^{-1} = \exp(-1) \approx 0.368$$

$$\left(1 + \frac{1}{n}\right)^n \xrightarrow[n \to \infty]{} e$$

So, $1 - e^{-1} \approx 0.632$ is the proportion (on average) of the data points that do end up in the bootstraped sample.

39. You are given a dataset with two variables, which is graphed below. You want to predict y using x.

Determine which statement regarding using a generalized linear model (GLM) or a random forest is true.



(A) A random forest is appropriate because the dataset contains only quantitative variables.

(B) A random forest is appropriate because the data does not follow a straight line.

(C) A GLM is not appropriate because the variance of y given x is not constant.

(D) A random forest is appropriate because there is a clear relationship between y and x.

(E) A GLM is appropriate because it can accommodate polynomial relationships.

Trees in general work well w/ qualitative predictors

The opposite is true.

The variance looks pretty constant

Could be anything!