(b) In most of the practical applications, not all the βs are important in modeling y, if we know the distribution instead of just the estimate itself of β, then we can delete some βs that are not significantly important under the significance level. which is very useful in variable selection in a regression model.

(c). $\because \hat{\beta} \sim N(\beta, \sigma^2(X^TX)^{-1})$

$\hat{y}_* = x_*^T \hat{\beta}$ also follows a Gaussian distribution.

$E(\hat{y}_*) = x_*^T \beta$

$Var(\hat{y}_*) = Var(x_*^T \hat{\beta}) = x_*^T [\sigma^2 (X^TX)^{-1}] x_*$

Therefore $\hat{y}_* \sim N(x_*^T \beta, \ \sigma^2 x_*^T (X^TX)^{-1} x_*$

(d). If we know the distribution of $\hat{y}_*$, we can compare it with the real data, if the two differs quite a lot, then this model would not be appropriate for modeling my data. Therefore knowing the distribution could be helpful to decide whether the model ~~is~~ justifiable to your real data.