

ECON_203_HW2_Ancel_Charles

Charles Ancel

2024-02-02

```
library(readxl)

# Load the dataset
housing_data <- read_excel("housing.xls")
```

Part1(a): Compute $P(P > 200|A = 3)$

To compute the probability of a house selling for more than \$200,000 given it has 3 bedrooms ($P(P > 200|A = 3)P(P > 200|A = 3)$), we first filtered the dataset for three-bedroom houses and then determined the proportion of these houses that sold for more than \$200,000.

```
num_houses_price_over_200_and_3bdrms <- sum(housing_data$price[housing_data$bdrms == 3] > 200)
total_3bdrm_houses <- sum(housing_data$bdrms == 3)
probability_over_200_given_3bdrms <- num_houses_price_over_200_and_3bdrms / total_3bdrm_houses

cat("Number of 3 bedroom houses:", total_3bdrm_houses, "\n")
```

```
## Number of 3 bedroom houses: 42
```

```
cat("Number of 3 bedrooms houses that sold for more than $200,000:", num_houses_price_over_200_and_3bdrms, "\n")
```

```
## Number of 3 bedrooms houses that sold for more than $200,000: 39
```

```
cat("Probability  $P(P > 200 | A = 3)$ :", probability_over_200_given_3bdrms, "\n")
```

```
## Probability  $P(P > 200 | A = 3)$ : 0.9285714
```

Results:

- Probability $P(P > 200|A = 3)$ is approximately 0.929

Part 1(b): Compute Probabilities for Different Prices and Number of Rooms

In this section, we extend our analysis to compute the conditional probabilities $P(P > p|A = a)$ for a range of prices p and bedroom counts a . Specifically, we look at bedroom counts of 2, 3, and 4 and prices ranging from \$100,000 to \$500,000 in increments of \$1,000.

```
price_range <- 100:500 # 100 to 500
bedroom_numbers <- c(2, 3, 4)

# Initialize a list to hold probabilities
probabilities <- list()

for (bdrms in bedroom_numbers) {
  prob_for_bdrms <- numeric(length(price_range))
```

```

for (i in 1:length(price_range)) {
  price <- price_range[i]
  num_houses_over_price <- sum(housing_data$price[housing_data$bdrms == bdrms] > price)
  total_houses <- sum(housing_data$bdrms == bdrms)
  prob_for_bdrms[i] <- num_houses_over_price / total_houses
}

probabilities[[as.character(bdrms)]] <- prob_for_bdrms
}

```

Part 1(c): Plotting Probabilities

To visualize how the probability of a house selling for more than a given price varies by the number of bedrooms, we plot $P(P > p | A = a)$ against price p for each bedroom count $a = 2, 3, 4$.

```

plot(price_range, probabilities[["2"]], type = "l", col = "red", ylim = c(0, 1), xlab = "Price ($1000s)")
lines(price_range, probabilities[["3"]], col = "blue")
lines(price_range, probabilities[["4"]], col = "green")
legend("topright", legend = c("2 bedrooms", "3 bedrooms", "4 bedrooms"), col = c("red", "blue", "green"))

```

$P(P > p | A = a)$ against p

