Extra practice with a current pop culture twist

Sci-fi coin flips

My partner and I are really enjoying the TV series Foundation (based on Isaac Asimov's book series).

Without giving away too many spoilers, there is a scene where a character correctly guesses the result of a coin flip 6 times in a row and says "Repeated luck is never luck". While as a statistician, I'd "never say never", but I think this is a cool way to demonstrate something about the character, namely that she has some sort of psychic ability.

Assume the coin is fair. Could this have just been luck?

```
store <- rep(NA, 1000)

for(i in 1:1000){
  sim <- sample(c(TRUE, FALSE), size = 6, replace=TRUE)

  prop <- mean(sim)

  store[i] <- prop
}
```

Question 1

What analysis is being conducted, above?

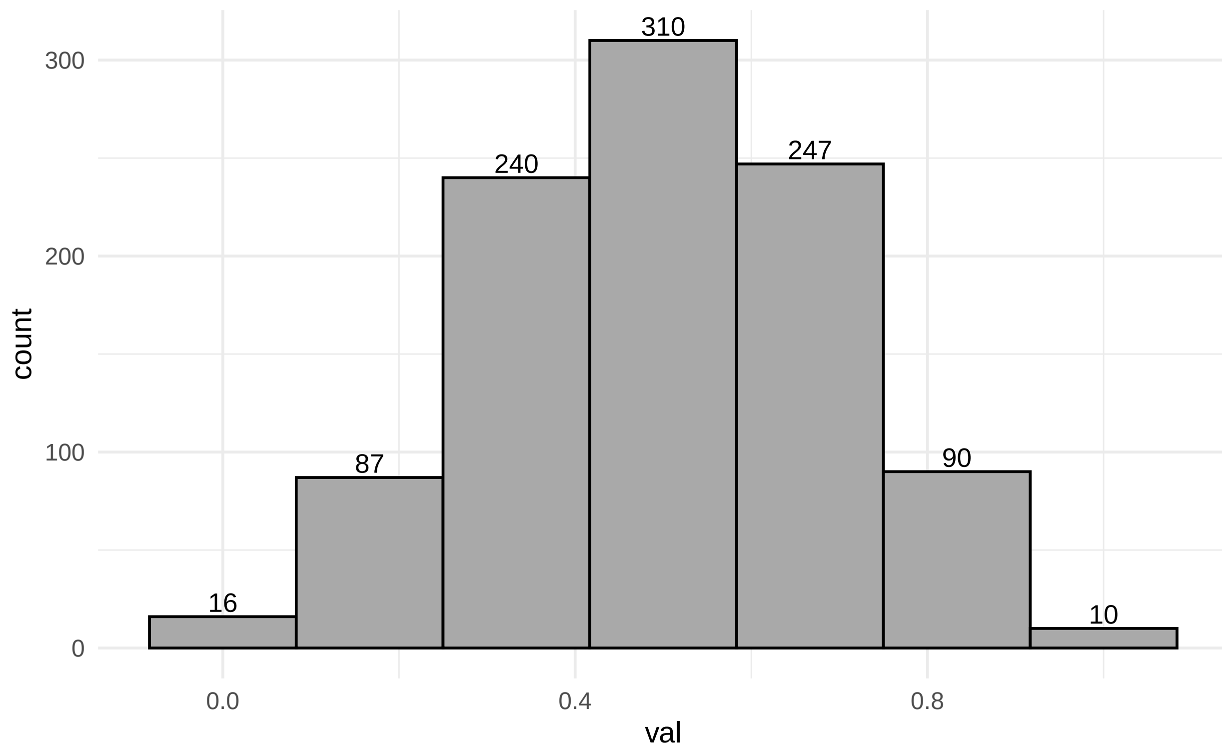
Question 2

What are then null and alternative hypothesis for this analysis?

Question 3

Calculate the p-value for this investigation, based on the information below

```
store_tibble_1 <- tibble(val = store)

store_tibble_1 %>%
  ggplot(aes(val)) +
  geom_histogram(bins = 7, color = "black", fill = "darkgrey") +
  theme_minimal() +
  stat_bin(bins=7, geom="text", colour="black", size=3.5,
           aes(label=..count..), vjust = -0.25)
```




```
store_tibble_1 %>%
  count(val)
```

```
## # A tibble: 7 x 2
##     val     n
##   <dbl> <int>
## 1 0        16
## 2 0.167    87
## 3 0.333   240
## 4 0.5     310
## 5 0.667   247
## 6 0.833    90
## 7 1        10
```

NOT the Squid Game

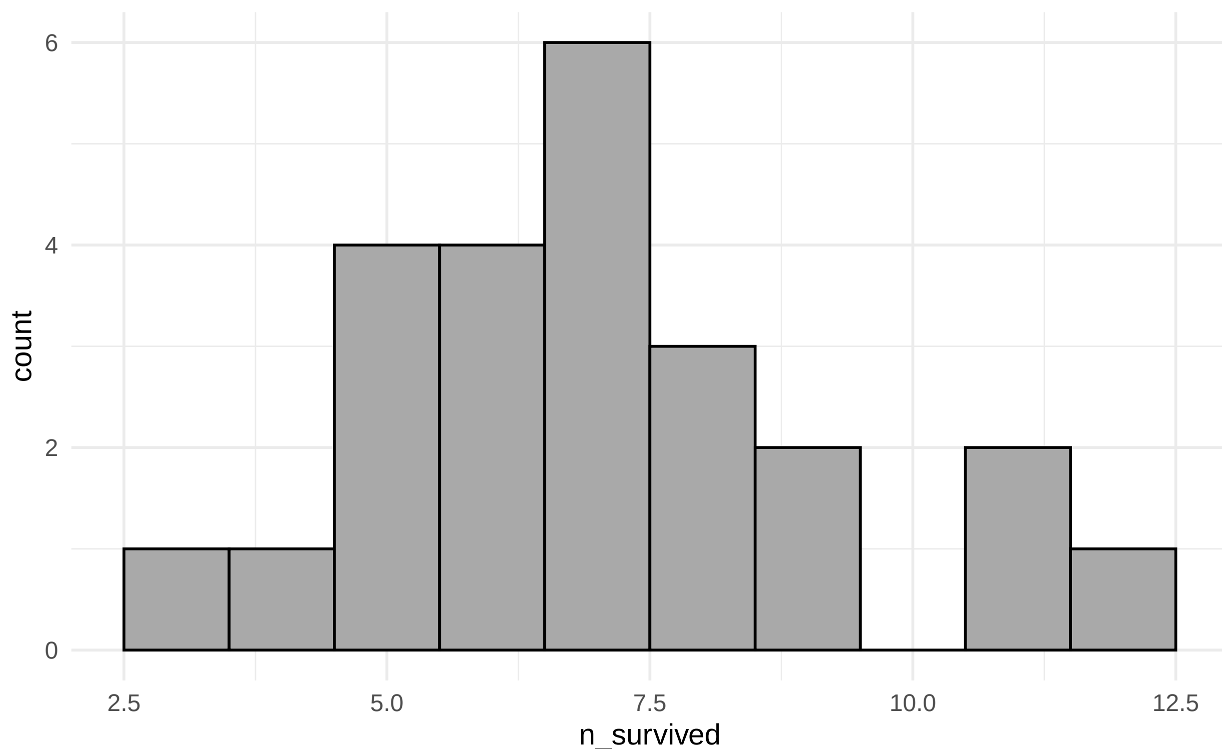
Suppose, for their own amusement, a secretive group of rich people create a series of games with life or death consequences for the players. Further suppose that one of these games involves a bridge of with 18 pairs of glass tiles. One tile of each pair can hold the weight of a person while the other cannot. Suppose, to test out the game concept, that the staff were asked to play a 'non-lethal' version of the game 24 times to get a grip on how many players were likely to survive. They played with 16 starting numbers of players. Assume that once one player made it across, all the players behind that play would also cross safely and there was no sabotaging of other players.

```
head(results)
```

```
## # A tibble: 6 x 3
```

```
##   first_player_across n_survived p_survived
##                 <int>      <dbl>      <dbl>
## 1                   8          9      0.562
## 2                  11          6      0.375
## 3                  10          7      0.438
## 4                  11          6      0.375
## 5                  12          5      0.312
## 6                  12          5      0.312
```

```
results %>%
  ggplot(aes(x = n_survived)) +
  geom_histogram(bins = 10, fill="darkgrey", color = "black") +
  theme_minimal()
```




```
results %>%
  summarise(x1 = min(n_survived),
            x2 = quantile(n_survived, 0.25),
            x3 = median(n_survived),
            x4 = quantile(n_survived, 0.75),
            x5 = max(n_survived),
            x6 = mean(n_survived),
            x7 = sd(n_survived))
```

```
## # A tibble: 1 x 7
##      x1    x2    x3    x4    x5    x6    x7
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1     3  5.75     7     8    12  7.04  2.22
```

Question 4

What was the minimum, maximum and average number of players that survived, in these tests?

Question 5

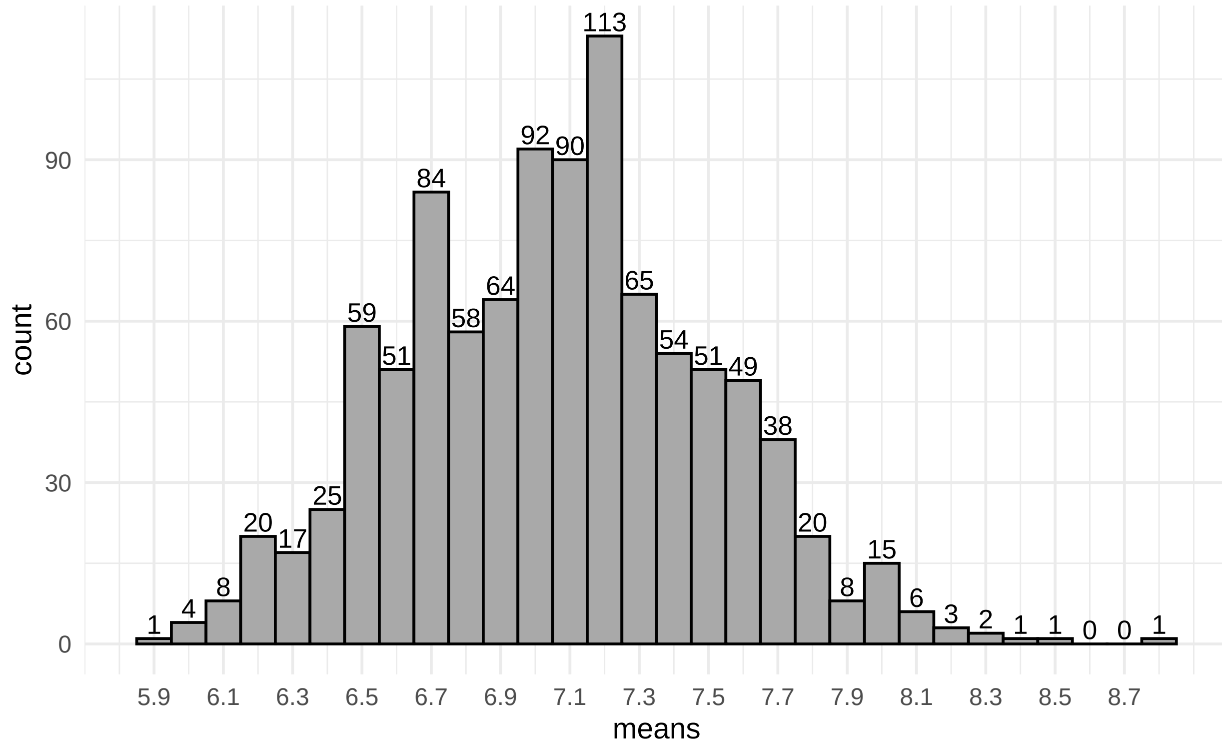
What analysis is being conducted in the chunk below?

```
set.seed(456)
store <- rep(NA, 1000)

for(i in 1:1000){
  boot_resample <- results %>%
    sample_n(24, replace=TRUE)

  store[i] <- mean(boot_resample$n_survived)
}
```

```
quantile(store, c(0.005, 0.025, 0.050, 0.1, 0.9, 0.950, 0.975, 0.995))
```

```
##     0.5%     2.5%       5%      10%      90%      95%    97.5%    99.5%
## 6.083125 6.208333 6.372917 6.500000 7.625000 7.791667 7.958333 8.250417
```

```
store_tibble_2 <- tibble(means = store)

store_tibble_2 %>%
  ggplot(aes(x = means)) +
  geom_histogram(binwidth = 0.1, color = "black", fill = "darkgrey") +
  theme_minimal() +
  stat_bin(binwidth = 0.1, geom="text", colour="black", size=3.5,
           aes(label=..count..), vjust = -0.25) +
  scale_x_continuous(breaks = round(seq(min(store), max(store), by = 0.2),1))
```



Question 6

Interpret the 95% confidence interval for this investigation.

Question 7

Suppose someone claimed they thought 9 people would survive on average, when 16 people played (i.e., 56.25%) and you were going to set up a hypothesis test using their claim as your null hypothesis value. What could you predict about your results for the hypothesis test with a significance cut-off of 5%, before conducting it, based on your 95% confidence interval?