Understanding Randomness

Chapter 9

Jason Bryer
epsy530.bryer.org

Random?

- Pick a number between 1 and 4. Is this random?
- How to pick a random number:
  - Pick a card from deck of cards.
  - Flip a coin.
  - Roll a dice.
  - Number out of a hat.
  - Use a computer.

```
sample(1:4, size = 1)
```

```
[1] 4
```

```
sample(1:4, size = 10, replace = TRUE)
```

```
 [1] 4 3 2 4 4 4 1 1 4 2
```

Simulation

- A cerial box contains one of three cards with the following chances: 20% for Walter White, 30% for Jesse Pinkman, or 50% for Hank Schrader.
- How many boxes would you expect to buy to get one of each?
- We can answer this using simulation.

Steps for Simulation

Specify how to model a component outcome using equally likely random digits:

1. Identify the component to be repeated: Opening a cereal box
2. Explain how you will model the experiment's outcome.
   Assign outcomes to the equally likely random digits 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 0 or 1 → Walt 2, 3, or 4 → Jesse 5, 6, 7, 8, or 9 → Hank

Specify how to simulate trials:

1. Explain how you will combine the components to model the trial. Look at each digit until all three types are found.
2. State clearly what the response variable is. We want to find the number of boxes it takes to get all three pictures.

Put it all together to run the simulation:

1. Run several trials

```
cards <- c(rep("Walt", 2), rep("Jesse", 3), rep("Hank", 5))
cards
```

```
 [1] "Walt"  "Walt"  "Jesse" "Jesse" "Jesse" "Hank"  "Hank"  "Hank"  "Hank"  "Hank"
```

Trial 1:

```
sample(cards, 8)
```

```
[1] "Hank"  "Jesse" "Hank"  "Hank"  "Jesse" "Hank"  "Hank"  "Walt"
```

Trial 2:

```
sample(cards, 8)
```

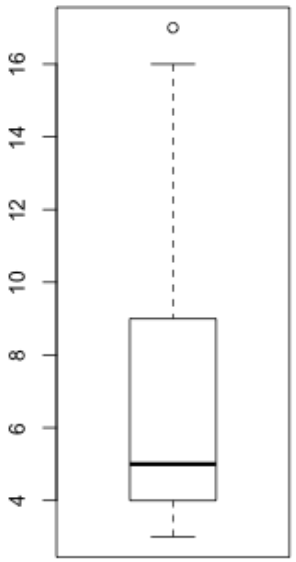
```
[1] "Hank"  "Jesse" "Walt"  "Hank"  "Walt"  "Hank"  "Jesse" "Hank"
```

Putting all together

```
nboxes <- integer()
for (i in 1:100) {
    test <- rep(FALSE, length(unique(cards)))
    names(test) <- unique(cards)
    counter <- 0
    while (any(!test)) {
        counter <- counter + 1
        test[sample(cards, 1)] <- TRUE
    }
    nboxes[i] <- counter
}
summary(nboxes)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   3.00    4.00    5.00    6.75    9.00   17.00
```

```
boxplot(nboxes)
```



More simulations

- Simulating coin flips and dice rolls.
  http://spark.rstudio.com/jbryer/gambler/
- Simulating lottery ticket purchaes.
  http://spark.rstudio.com/jbryer/gambler/

Summary

Don’t Overstate Your Case

* Simulation is not reality, it only indicates probability.

Model Outcome Chances Accurately

* What would be wrong with generating random numbers 0, 1, 2, 3 to indicate the number of team members? * There is not a 25% chance of each. They are not equally likely.

Run Enough Trials

* Don’t just do a few trials. Err on the side of a large number of trials.