Lab 9a: Expected Value and Std Error

Stat 131A, Fall 2018

Learning Objectives:

- Expected Value of Common Random Variables.
- Standard Error of Common Random Variables.

General Instructions

- Write your solutions in an `Rmd` (R markdown) file.
- Name this file as `lab09a-first-last.Rmd`, where `first` and `last` are your first and last names (e.g. `lab09a-gaston-sanchez.Rmd`).
- Knit your `Rmd` file as an html document (default option).
- Submit your `Rmd` and `html` files to bCourses, in the corresponding lab assignment.

Problem 1

One hundred draws are made at random with replacement from the box `[1, 2]`.

a. How small can the sum be?

b. How large can the sum be?

c. How many times do you expect the ticket `1` to turn up?

d. How many times do you expect the ticket `2` to turn up?

e. About how much do you expect the sum to be?

To help you understand this chance process, you can use the following R code that simulates repeating the process 1000 times (each time 100 tickets are drawn out of the box). And then plotting a histogram for the sum of the tickets.

```
# box
box = c(rep(1, 50), rep(2, 50))

# 1000 repetitions each of which involves drawing 100 tickets
# (in each repetition the sum of tickets is calculated)
repetitions = 1000
sum_tickets = rep(0, repetitions)
```

```r
set.seed(12017)
for (i in 1:repetitions) {
  # 100 draws
  draws = sample(box, size = 100, replace = TRUE)
  # sum of tickets
  sum_tickets[i] = sum(draws)
}

# histogram for sum of tickets
hist(sum_tickets, breaks = 30, col = 'gray80', las = 1)
```

Problem 2

Find the expected value for the sum of 100 draws at random with replacement from each of the following boxes:

a. `[0, 1, 1, 6]`

b. `[-2, -1, 0, 2]`

c. `[-2, -1, 3]`

d. `[0, 1, 1]`

Problem 3

Fifty draws are made at random with replacement from the box `[1 2 3 4 5]`. The sum of the draws turns out to be 157.

a. The expected value for the sum is _____.

b. The standard error for the sum is _____.

Problem 4

Tickets are drawn at random with replacement from a box of numbered tickets. The sum of 25 draws has expected value equal to 50, and the SE is 10. If possible, find the expected value and SE of the sum of 100 draws. Or do you need more information?

Problem 5

Suppose a computer is programmed to take the sum of 25 draws made at random with replacement from the box: `[0 2 3 4 6]`. The program prints out the result, repeating the process over and over again.

About what percentage of the observed values should be between 50 and 100?

Problem 6

Consider the box `[1 1 2 2 2 4]`. One hundred draws will be made at random with replacement from this box.

a. The smallest the sum can be is _____

b. The largest the sum can be is _____

c. The sum of the draws will be around _____, give or take _____ or so.

d. The chance that the sum will be bigger than 250 is almost _____.

Problem 7

Consider the box `[-1 1]`. You can draw either 10 times or 100 times at random with replacement from this box. How many times should you draw:

a. To win $1 when the sum is 5 or more, and nothing otherwise?

b. To win $1 when the sum is -5 or less, and nothing otherwise?

c. To win $1 when the sum is between -5 and 5, and nothing otherwise?

Problem 8

There are two options:

(i) One hundred independent draws from the box: `[1 1 5 7 8 8]`

(ii) Twenty-five independent draws from the box: `[14 17 21 23 25]`

Which is better, if the payoff is:

a. $1 when the sum is 550 or more, and nothing otherwise?

b. $1 when the sum is 450 or more, and nothing otherwise?

c. $1 when the sum is between 450 and 550, and nothing otherwise?

Problem 9

A biased coin has one chance in ten of landing heads. It is tossed 400 times. Use the normal approximation, with continuity correction, to estimate the chance of getting exactly 40 heads. Compare your answer with the one produced by `dbinom()`.

Problem 10

Consider the box: `[0 1 2 3]`. Ten draws are going to be made at random with replacement from this box. Use the normal approximation, with continuity correction, to approximate the chance that the sum will be in the interval from 10 to 20 inclusive.

Problem 11

A coin is tossed 100 times.

a. The difference *"number of heads - number of tails"* is like the sum of 100 draws from one of the following boxes. Which one, and why?

```
(i)    [heads, tails]
(ii)   [-1, 1]
(iii)  [-1, 0]
(iv)   [0, 1]
(v)    [-1, 0, 1]
```

b. Find the expected value and standard error for the difference. Show your work.

Problem 12

A gambler plays roulette 100 times, betting a dollar on a column each time. The bet pays 2 to 1, and there are 12 chances in 38 to win. Fill in the blanks.

a. In 100 plays, the gambler's net gain will be aorund $_____, give or take $_____ or so.

b. In 100 plays, the gambler should win _____ times, give or take _____ or so.

Problem 13

A box contains 10 tickets. Each ticket is marked with a whole number between -4 and 4. The numbers are not all the same; their average equals 0. You have two choices:

i. 50 draws are made from the box, and you win $10 if the sum is between -12 and 12.

ii. 100 draws are made from the box, and you win $10 if the sum is between -24 and 24.

Choose one of the four options below; explain your reasoning.

a. (i) and (ii) give the same chance of winning.

b. (i) gives a better chance of winning.

c. (ii) gives a better chance of winning.

d. Can't tell without the SD of the box.

To help you understand this chance process, you can use the following R code that simulates repeating each process 10000 times (in one of them 50 tickets are drawn, in the other 100 tickets are drawn). And then plotting the frequency for the sums of the tickets.

```
# box
box = rep(c(-4, 4), 5)

# 10000 repetitions each of which involves drawing 50 tickets
# (in each repetition the sum of tickets is calculated)
repetitions = 10000
sum_tickets50 = rep(0, repetitions)

set.seed(12017)
for (i in 1:repetitions) {
  # 50 draws
  draws50 = sample(box, size = 50, replace = TRUE)
  # sum of tickets
  sum_tickets50[i] = sum(draws50)
}

# freq. histogram for sum of tickets
barplot(table(sum_tickets50), border = NA, las = 1)


# 10000 repetitions each of which involves drawing 100 tickets
# (in each repetition the sum of tickets is calculated)
repetitions = 10000
sum_tickets100 = rep(0, repetitions)

set.seed(12017)
for (i in 1:repetitions) {
  # 50 draws
  draws100 = sample(box, size = 100, replace = TRUE)
  # sum of tickets
  sum_tickets100[i] = sum(draws100)
}

# freq. histogram for sum of tickets
barplot(table(sum_tickets100), border = NA, las = 1)
```