Exercise 1

Spring 2021 STAT400 Homework 7 Solutions

Yu take a trip to Curtis Orchard are interested in Ambrosia and Winesap apples. Assume the following:

- All apple weights are independent.
- The weight of the Ambrosia apples is normally distributed with a mean of 90 grams and a standard deviation of 4 grams. Let A be the weight of a randomly selected Ambrosia apple.
- The weight of the Winesap apples is normally distributed with a mean of 88 grams and a standard deviation of 6 grams. Let W be the weight of a randomly selected Winesap apple.

a) (0.5 pt) Suppose you pick 5 Winesap apples at random. Assuming independence, what is the probability that that the **average** weight of the 5 applies is less than 89 grams?

b) (0.5 pt) Suppose you pick 5 Ambrosia apples at random. Assuming independence, what is the probability that that the **total** weight of the 5 applies is more than 446 grams?

c) (0.5 pt) Suppose you pick one Ambrosia and One Winesap apple at random. What is the probability that the Ambrosia apple weighs less than the Winesap apple?

d) (1 pt) Suppose you pick 5 Winesap apples and 5 Ambrosia apples. What is the probability that their total weight is less than 900g?

e) (0.5 pt) Suppose you continue to pick Winesap apples until you get three that weigh over 95g. Let X represent the number of apples you must pick. Find E[X].

$A \sim \text{Normal}(\mu = 90, \sigma^2 = 16)$; $W \sim \text{Normal}(\mu = 88, \sigma^2 = 36)$

from calculator/R

from Normal Table

a) $\bar{W} \sim \text{Normal}(\mu = 88, \sigma^2 = 36/5)$

$$P[\bar{W} < 89] = P\left[Z < \frac{89-88}{6/\sqrt{5}}\right] \approx 0.6453$$

$$\approx P[Z < 0.37] \approx 0.6443$$

b) $\sum_{i=1}^{5} A_i \sim \text{Normal}(\mu = 5 \cdot 90, \sigma^2 = 5 \cdot 16)$

$$P\left[\sum_{i=1}^{5} A_i > 446\right] = P\left[Z > \frac{446-450}{4\sqrt{5}}\right] \approx 0.6726$$

$$\approx P[Z > -0.45] \approx 0.6736$$

c) Solve $P[A < W] \rightarrow P[A - W < 0]$

Let X be the distribution of $A - W$: $X \sim \text{Normal}(\mu = 90 - 88, \sigma^2 = 16 + 36)$

$$P[X < 0] = P\left[Z < \frac{0-2}{\sqrt{52}}\right] \approx 0.3908$$

$$\approx P[Z < -0.28] \approx 0.3897$$

d) Solve $P\left[\sum_{i=1}^{5} W_i + A_i < 900\right]$

Let Y be $\sum_{i=1}^{5} W_i + A_i \rightarrow Y \sim \text{Normal}(5(90+88), 5(16+36))$

$$P[Y < 900] = P\left[Z < \frac{900-890}{2\sqrt{65}}\right] \approx 0.7324$$

$$\approx P[Z < 0.62] \approx 0.7324$$

e) $P[W > 95] = P\left[Z > \frac{95-88}{6}\right] = P\left[Z > \frac{7}{6}\right] \approx 0.1218^{*}$ ← probability of one Winesap Apple

$X \sim \text{NBinomial}(r = 3, p = 0.1218) \rightarrow \mu = E[X] = r\left(\frac{1}{p}\right) \approx 24.66$ apples

$= 3\left(\frac{1}{0.121}\right) \approx 24.79$ apples

Exercise 2

Consider two random variables X and Y, where

- $\sigma_X = 5$
- $\sigma_Y = 2$
- $Var[2X - 3Y] = 80$

(2 pts) Calculate the correlation between X and Y, ρ_{XY}.

$$\sigma_x = 5 \rightarrow \sigma_x^2 = 25$$
$$\sigma_y = 2 \rightarrow \sigma_y^2 = 4$$

$$Var[2X-3Y] = (2^2)\sigma_x^2 - (2)(6)\sigma_{xy} + (3^2)\sigma_y^2$$
$$80 = 4(25) - 12\sigma_{xy} + 9(4)$$
$$12\sigma_{xy} = 56$$
$$\sigma_{xy} = 14/3 \approx 4.667$$
$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{14/3}{5 \cdot 2} = 7/15 \approx 0.4667$$

Exercise 3

Let $X_1, X_1, ..., X_{100}$ be a i.i.d. random sample of size n=100 from a distribution with probability density function:

$$f(x) = 6x(1-x),\ 0 < x < 1$$

(3 pts) Approximate

$$P(0.45 < \bar{X} < 0.5), \qquad where\ \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Hint: find μ and σ^2.

$$\mu = E[X] = \int_0^1 6x^2 - 6x^3\,dx = \left[2x^3 - \tfrac{3}{2}x^4\right]_0^1 = 1/2 = 0.50$$
$$E[X^2] = \int_0^1 6x^3 - 6x^4\,dx = \left[\tfrac{3}{2}x^4 - \tfrac{6}{5}x^5\right]_0^1 = 3/10 = 0.30$$
$$\sigma^2 = Var[X] = E[X^2] - (E[X])^2 = \tfrac{3}{10} - \left(\tfrac{1}{2}\right)^2 = 1/20 = 0.05$$
$$\bar{X} \sim Normal\left(\mu = \tfrac{1}{2}, \sigma^2 = \tfrac{1}{2000}\right)$$
$$P[0.45 < \bar{X} < 0.50] = P[\bar{X} < 0.50] - P[\bar{X} < 0.45] = P\left[Z < \frac{0.50-0.50}{1/\sqrt{2000}}\right] - P\left[\frac{0.45-0.50}{1/\sqrt{2000}}\right]$$
$$= P[Z<0] - P[Z<-\sqrt{5}] \approx 0.50 - 0.0128 \approx 0.4873$$
$$\approx P[Z<0] - P[Z<-2.24] \approx 0.50 - 0.0125 \approx 0.4875$$

Exercise 4

Because of the randomness, the numbers are not important; however, the code must follow similar logic as the three variants below.

Let A ~ N(90, 16) and W ~ N(88, 36).

a) (1 pt) Use R to generate 2 independent, random samples of size 10 from each of these distributions (A and W). Comparing by elements, what proportion of A are smaller than W?

Defining a function (not necessary but convenient):

```
manzanas = function(sample_size) {
  ambrosias = rnorm(n = sample_size, mean = 90, sd = 4)
  winesaps = rnorm(n = sample_size, mean = 88, sd = 6)
  return(mean(ambrosias < winesaps))
}
```

n = 10

```
prop_10 = manzanas(sample_size = 10)
prop_10
```

```
## [1] 0.1
```

The proportion of A that is smaller than W is 0.1.

b) (1 pt) Repeat 4(a) with samples of size n = 100 and n = 10000. What proportion of A are elementwise smaller than W? (Compare your answer to 1.c))

n = 100

```
prop_100 = manzanas(sample_size = 100)
prop_100
```

```
## [1] 0.4
```

The proportion of A that is smaller than W in this random sample of size n = 100 is 0.4.

n = 10000

```
prop_10000 = manzanas(sample_size = 10000)
prop_10000
```

```
## [1] 0.3952
```

The proportion of A that is smaller than W in this random sample of size n = 10000 is 0.3952.

Notice that, as n increases, the proportion of elementwise A < W generally becomes closer and closer to the answer from Exercise 1 Part c, which is the true proportion/expectation of the population. This is because a large sample has a narrower sampling distribution than a smaller one (in other words, higher n leads to lower variance), therefore coming closer to the true proportion/expectation.

Note: Other successful methods include sum(booleans)/sample_size, for loops, and more, but we recommend this method for its concision.