STAT5002 Weekly Independent Exercises

Sheet 3 - Week 6

STAT5002

1 0-1 Box (specific example)

A box contains 10 tickets. 3 are $\boxed{1}$ and 7 are $\boxed{0}$. In the questions below, if necessary, round to 3 decimal places.

1.1

What is the mean μ and SD σ of the box (that is, of the list of numbers represented on the tickets in the box)?

1.2

Suppose $n = 100$ tickets are drawn randomly, with replacement, yielding numbers $X_1, \ldots, X_n$. Write $S = X_1 + \ldots + X_n$ for the sum of the draws and $\bar{X} = S/n$ for the average of the draws.

1.2.1

What is $E(X_1)$?

1.2.2

What is $SE(X_1)$?

1.2.3

What is $E(X_1 + X_2)$?

1.2.4

What is $SE(X_1 + X_2)$?

1.2.5

What is $E(S)$?

1.2.6

What is $SE(S)$?

1.2.7

What is $E(\bar{X})$?

1.2.8

What is $SE(\bar{X})$?

1.3

By appealling to the Central Limit Theorem, determine a value v such that the interval $0.3 \pm v$, i.e. $[0.3 - v, 0.3 + v]$, serves as an (approximate) 98% prediction interval for $\bar{X}$. In other words, find v such that

$$P\left\{0.3 - v \leq \bar{X} \leq 0.3 + v\right\} \approx 0.98\,.$$

The R output below may be useful for this.

```
qnorm(c(0.95, 0.975, 0.98, 0.99))
```

```
[1] 1.644854 1.959964 2.053749 2.326348
```

2 0-1 Box (general case)

Repeat question 1, but for a box with N tickets: pN are $\boxed{1}$ and $(1-p)N$ are $\boxed{0}$. Write out answers to the questions below in terms of general sample size n and proportion of $\boxed{1}$s p.

2.1

What is the mean μ and SD σ of the box (that is, of the list of numbers represented on the tickets in the box)?

2.2

Suppose n tickets are drawn randomly, with replacement, yielding numbers $X_1, \ldots, X_n$. Write $S = X_1 + \ldots + X_n$ for the sum of the draws and $\bar{X} = S/n$ for the average of the draws. You may assume that n is large enough that the Central Limit Theorem applies.

2.2.1

What is $E(X_1)$?

2.2.2

What is $SE(X_1)$?

2.2.3

What is $E(X_1 + X_2)$?

2.2.4

What is $SE(X_1 + X_2)$?

2.2.5

What is $E(S)$?

2.2.6

What is $SE(S)$?

2.2.7

What is $E(\bar{X})$?

2.2.8

What is $SE(\bar{X})$?

2.3

By appealling to the Central Limit Theorem, determine a value v such that the interval $p \pm v$, i.e. $[p - v, p + v]$, serves as an (approximate) 98% prediction interval for $\bar{X}$. In other words, find v such that

$$P\left\{p - v \leq \bar{X} \leq p + v\right\} \approx 0.98\,.$$

The R output below question 1.3 may be useful for this.