CS 361 Spring 2018
Homework 6

Nathaniel Murphy (njmurph3)

**Note that in this problem set we will be using X to denote the population of what we are sampling.

6.2

(a)

The best estimate for population mean of the weight of the mice is mean of the sample of the mice.

$$\text{popmean}(\{X\}) = \text{mean}(\{x\}) = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{199}{10} = \mathbf{19.9} \text{ grams}$$

(b)

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}}, \; \text{stdunbiased}(\{x\}) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \text{mean}(\{x\}))^2}{N-1}}$$

$$\text{stdunbiased}(\{x\}) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \text{mean}(\{x\}))^2}{N-1}} = \sqrt{\frac{110.9}{9}} = 3.51$$

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = \frac{3.51}{\sqrt{10}} = \mathbf{1.11} \text{ grams}$$

(c)

Assuming that stdunbiased($\{x\}$) parameter does not change as the sample size N changes, we have the equation:

$$\text{std}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = 0.1$$

$$\frac{\text{stdunbiased}(\{x\})}{0.1} = \sqrt{N}$$

$$\left(\frac{\text{stdunbiased}(\{x\})}{0.1}\right)^2 = N$$

$$\left(\frac{3.51}{0.1}\right)^2 = \mathbf{1233} \text{ mice}$$

6.3

Let us assume that we have the relation yellow $\to 0$ and blue $\to 1$.

(a)

$$\text{stdunbiased}(\{x\}) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \text{mean}(\{x\}))^2}{N-1}} = \sqrt{\frac{7(-0.3)^2 + 3(0.7)^2}{9}} = 0.483$$

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = \frac{0.483}{\sqrt{10}} = \mathbf{0.153}$$

(b)

Assuming that stdunbiased($\{x\}$) does not change as N changes, we have the algorithm:

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = 0.05$$

$$\frac{\text{stdunbiased}(\{x\})}{0.05} = \sqrt{N}$$

$$\left(\frac{\text{stdunbiased}(\{x\})}{0.05}\right)^2 = N$$

$$\left(\frac{0.483}{0.05}\right)^2 = \mathbf{94} \text{ draws}$$

6.4

Given $N = 40$, stdunbiased($\{x\}$) $= 75$ grams, let us compute stderr($\{x\}$).

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = \frac{75}{\sqrt{40}} = 11.859 \text{ grams}$$

(a)

A 68% confidence interval for popmean($\{x\}$) is given by the following formula:

$$\text{mean}(\{x\}) - \text{stderr}(\{x\}) \leq \text{popmean}(\{x\}) \leq \text{mean}(\{x\}) + \text{stderr}(\{x\})$$

$$340 - 11.859 \leq \text{popmean}(\{x\}) \leq 340 + 11.859$$

$$328.141 \leq \text{popmean}(\{x\}) \leq 351.859$$

(b)

A 99% confidence interval for popmean($\{x\}$) is given by the following formula:

$$\text{mean}(\{x\}) - 3 \cdot \text{stderr}(\{x\}) \leq \text{popmean}(\{x\}) \leq \text{mean}(\{x\}) + 3 \cdot \text{stderr}(\{x\})$$

$$340 - 3(11.859) \leq \text{popmean}(\{x\}) \leq 340 + 3(11.859)$$

$$304.123 \leq \text{popmean}(\{x\}) \leq 375.577$$

6.5

We observe that z-values for 0.8 and 0.95 are 0.845 adn 1.96, respectively.

The sample dataset is identical to that in problem 6.2, so we may say that mean($\{x\}$) = 19.9 grams and stderr($\{x\}$) = 1.11 grams

(a)

The 80% confidence interval for popmean($\{X\}$):

$$\text{mean}(\{x\}) - 0.845(\text{stderr}(\{x\})) \leq \text{popmean}(\{x\}) \leq \text{mean}(\{x\}) + 0.845(\text{stderr}(\{x\}))$$

$$19.9 - 0.845(1.11) \leq \text{popmean}(\{x\}) \leq 19.9 + 0.845(1.11)$$

$$18.962 \leq \text{popmean}(\{x\}) \leq 20.838$$

(b)

The 95% confidence interval for popmean($\{x\}$):

$$\text{mean}(\{x\}) - 1.96(\text{stderr}(\{x\})) \leq \text{popmean}(\{x\}) \leq \text{mean}(\{x\}) + 1.96(\text{stderr}(\{x\}))$$

$$19.9 - 1.96(1.11) \leq \text{popmean}(\{x\}) \leq 19.9 + 1.96(1.11)$$

$$17.724 \leq \text{popmean}(\{x\}) \leq 22.076$$

6.6

Let X be the random variable of a child that is born in Carcelle-le-Grignon. Let x denote the random sample. Assuming that female $\rightarrow$ 1 and male $\rightarrow$ 0, we have that

- $P(x = 0) = \frac{983}{2009} = 0.4892$
- $P(x = 1) = \frac{1026}{2009} = 0.511$

$$\text{stdunbiased}(\{x\}) = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \text{mean}(\{x\}))^2}{N-1}} = \sqrt{\frac{983(0.511)^2 + 1026(0.4892)^2}{2008}} = 0.5005$$

$$\text{stderr}(\{x\}) = \frac{\text{stdunbiased}(\{x\})}{\sqrt{N}} = \frac{0.5005}{\sqrt{2009}} = 0.0112$$

(a)

The 99% confidence interval for $P(X = 1)$:

$$P(x = 1) - 3 \cdot \text{stderr}(\{x\}) \leq P(X = 1) \leq P(x = 1) + 3 \cdot \text{stderr}(\{x\})$$

$$0.511 - 3(0.0112) \leq P(X = 1) \leq 0.0511 + 3(0.0112)$$

$$0.4771 \leq P(X = 1) \leq 0.5443$$

(b)

The 99% confidence interval for $P(X=0)$:

$$P(x=0) - 3 \cdot \text{stderr}(\{x\}) \leq P(X=0) \leq P(x=0) + 3 \cdot \text{stderr}(\{x\})$$

$$0.4892 - 3(0.0112) \leq P(X=0) \leq 0.4892 + 3(0.0112)$$

$$0.4557 \leq P(X=0) \leq 0.5229$$

(c)

These intervals do overlap, so this implies that the probabilities could be equal of having a boy or having a girl.