HW 1

1a) There are 400 possible choices for the first sample, 399 for the second and so on

$$400 \cdot 399 \cdot \ldots \cdot 301 = \binom{400}{300}$$

1b) When sampling with replacement, the population does not reduce each time. It is not possible to write this in $\binom{n}{r}$ format.

$$n^r = 400^{100}$$

1c) Sample 100 with replacement from $\{1, 2, \ldots, 400\}$

The Central Limit Theorm states that the average of independent identically distributed samples will follow a normal distribution

Theoretical Parameters

$$\mu = \sum_{i=1}^{400} i/400 = 200.5 \cdot \frac{100}{100} = 200.5$$

$$\sigma = \sqrt{\sum_{i=1}^{400} \frac{i^2}{400} - \mu^2} = 115.47 \cdot \frac{\sqrt{100}}{100} = 11.54$$

11.76

This term is applied because we have 100 samples.

2a)
$$\begin{bmatrix} -0.1 & -1.2 & 0.1 \\ 1.5 & -0.5 & 0.5 \\ -0.6 & 1.0 & -0.5 \\ -1.4 & -1.5 & -2.0 \end{bmatrix}$$

2b)
$$B = \begin{bmatrix} 4.58 & 0.87 & 3.84 \\ 0.87 & 4.94 & 2.13 \\ 3.84 & 2.13 & 4.51 \end{bmatrix} \quad C = \begin{bmatrix} 1.46 & 0.50 & -1.19 & 1.74 \\ 0.50 & 2.75 & -1.65 & -2.35 \\ -1.19 & -1.65 & 1.61 & 0.34 \\ 1.74 & -2.35 & 0.34 & 8.21 \end{bmatrix}$$

2d) solve (c) failed. The reason stated by R is that ~~x~~ is singular. Indeed, checking det (c) = -5.6 e -16 a very small, nearly zero number.
ginv (C) did not pick up on this. I won't use that function in the future.

2c)
$$B^{-1} = \begin{bmatrix} 0.922 & 0.221 & -0.890 \\ 0.221 & 0.307 & -0.333 \\ -0.890 & -0.333 & 1.13 \end{bmatrix} \quad C^{-1} = \begin{bmatrix} 0.743 & -0.736 & -0.417 & -0.344 \\ -0.736 & 1.845 & 0.318 & 0.445 \\ -0.417 & 0.318 & 0.265 & 0.176 \\ -0.344 & 0.445 & 0.176 & 0.314 \end{bmatrix}$$

2e) row.median = $\begin{bmatrix} -0.35 \\ -0.85 \\ -0.20 \end{bmatrix}$ col.std = [0.700 1.000 0.896 0.321]

3a) length(y) = 62
sum(!is.na(y)) = 48
They are different b/c there are 14 nan values in the NonD column.

3b) w = na.omit(y)
w = y[!is.na(y)]

3c)

BodyWgt	BrainWgt	NonD	Dream	Sleep	Span	Gest
100.81	218.68	8.74	1.90	10.64	19.37	129.94

3d) reference R code

3e) [13.04 11.75 10.31 8.81 4.07]

4d) mean = 349.96
Var = 291.26

4e) View R code for histogram

4f) Normal distribution. There is a uniform random event, dice roll, and X is the sum of 100 dice rolls. The central limit theorem states that the sum (or average) of many independent, identically distributed values will be normal.

$Mean(\mu) = (1+2+3+4+5+6)/6 \cdot 100 = 350$
$Variance(\sigma^2) = [(1^2+2^2+3^2+4^2+5^2+6^2)/6 - \mu^2] \cdot 100 = 291.67$

These are similar to what I found experimentally in (d)